

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 14:06:48 ; Search time 22.0758 Seconds
(without alignments)
1638.789 Million cell updates/sec

Title: US-10-672-399-4
Perfect score: 1987
Sequence: 1 MRQQDAPKPTPAARCSGLA.....TRRWSPACSSWRPLCCVC 376
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1426	71.8	543	JC4812	hyaluronan synthase
2	802.5	40.4	588	A43740	DG42 protein - Afr
3	308	15.5	568	T17588	hyaluronoglucosami
4	213	10.7	424	Z2ZRCL	nodulation protein
5	208.5	10.5	426	A95321	NODC N-ACETYLGLUCO
6	205.5	10.3	426	Z2ZRC4	nodulation protein
7	199.5	10.0	413	S34305	nodulation protein
8	176.5	8.9	424	S12793	nodulation protein
9	163.5	8.2	1086	JG6079	chitin synthase (E
10	159.5	8.0	395	JQ0396	nodulation protein
11	159	8.0	1195	S61886	chitin synthase (E
12	156.5	7.9	1165	S45879	chitin synthase (E
13	153.5	7.7	1239	T42020	class IV chitin sy
14	151.5	7.6	1112	T30202	probable chitin sy
15	148	7.4	419	A33100	hyaluronate syntha
16	146	7.3	395	A48755	hyaluronan synthase
17	145.5	7.3	1041	T31097	chitin synthase (E
18	142	7.1	1175	S39951	chitin synthase (E
19	140	7.0	743	T34632	probable bi-functi
20	137.5	6.9	428	E38180	nodulation protein
21	135.5	6.7	366	E59102	hypothetical prote
22	123.5	6.2	1103	T42022	probable chitin sy
23	119.5	6.0	1198	T28678	polyketide synthase
24	115.5	5.8	447	A97211	glycosyltransferase
25	114.5	5.8	1498	S78102	chitin synthase (E
26	114.5	5.8	1852	JCS546	chitin synthase (E
27	114.5	5.8	1869	A59290	class V chitin syn
28	111.5	5.6	412	B90075	intercellular adhe
29	111	5.6	534	T35190	hypothetical prote

30 111 5.6 676 1 WMBEX6 UL6 protein - huma
31 108.5 5.5 869 2 H83500 probable glucosyl
32 106 5.3 487 2 S61243 deoxyribonuclease
33 106 5.3 3519 2 S43048 polyketide synthase
34 105 5.3 615 2 E70663 probable ppg prote
35 103 5.2 3573 2 S23070 erythronolide synt
36 100.5 5.1 3034 2 T14119 seven-pass transme
37 100 5.0 1155 2 AC2675 chromosome segrega
38 100 5.0 1165 2 A97457 structural mainten
39 99.5 5.0 291 2 B82645 phenylacetaldehyde
40 99.5 5.0 412 2 S77608 probable intercell
41 99.5 5.0 515 2 AC2635 glycosyltransferas
42 99.5 5.0 645 2 B97417 probable membrane
43 99 5.0 352 2 F75099 rhamnosyl transfer
44 99 5.0 6420 2 T30283 polyketide synthase
45 98.5 5.0 2314 2 T28698 hypothetical prote

ALIGNMENTS

RESULT 1

JC4812
hyaluronan synthase (EC 2.4.1.1) - human
C:Species: Homo sapiens (man)
C:Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 02-Mar-2001
C:Accession: JC4812
R:Itano, N.; Kimata, K.
Biochem. Biophys. Res. Commun. 222, 816-820, 1996
A:Title: Molecular cloning of human hyaluronan synthase.
A:Reference number: JC4812; MUID:96244584; PMID:8651928
A:Accession: JC4812
A:Molecule type: mRNA
A:Residues: 1-543 <ITA>
A:Cross-references: DDBJ:D84424; NID:gl401033; PIDN:BAAL2351.1; PID:d1013030; PID:gl401030
C:Comment: This enzyme synthesizes a high molecular weight glycosaminoglycan composed of
C:Genetics:
A:Gene: GDB:HAS1; HAS
A:Cross-references: GDB:L220109; OMIM:601463
A:Map position: 19q13.4-19q13.4
C:Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein
F:58,167,308,321/Binding site: phosphate (Thr) (covalent) #status predicted
F:82,247/Binding site: phosphate (Ser) (covalent) #status predicted

HA1a

RESULT 2

A43740
DG42 protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)

Query Match	10.3%;	Score	205.5;	DB 1;	Length	426;			
Best Local Similarity	25.7%;	Pred. No.	1.2e-08;						
Matches	80;	Conservative	31;	Mismatches	101;	Indels	99;	Gaps	10;

QY	49	DRYGLAFLGGLYGAFLSAHLVAQSLFAYLHRRV-AAAARGPLDAATARSVALTISAYQE	106
DB	5	DTTSTAISITAYULLTAYRSMQVLYA----RPIDGPAVSAEPVETRLPDAVDIVPFSNE	60
QY	107	DPAYLROCLASARALLYPRARLRVLMVVDGNRAEDLYMVMDFREVFADEDPATYVMDGNY	166
DB	61	DPGILSACLASIADQYP-GELRVYVDDGSRNR-----	93
QY	167	HQPWEPAAGAGVAGAYREVEAEDPGRLAVEALVTRRCVCVCAQRMG-----GKREV	218
DB	94	-----EAVRVAFYSRDPREFSFILLPENVGKKA	123
QY	219	MYTAFKALGDSV-DYVQVCDSDTRLDPMALLELVRVLDEDPRGVAGGVDRIINPLDSDW	277
DB	124	Q-----IAAIGSSGDLVNLVDSDTIAFDVVVSKLAKM-RDPEVGVVGGQLTASNSGDTWL	179
QY	278	SFLSSRLYVAFNVNERACQSYFHCVSCISGSLGTPPGPAATQRRP-----	322
DB	180	TKLIDMEYLACNEERRAAQSRFGAVMCCC-----GPCMYRRSALASLLDQYETOLFR	232
QY	323	-RPSCGGASRH	332
DB	233	GKPSDFGEDRH	243

RESULT 7

S34305

modulation protein nodC - Rhizobium sp.

C:Species: Rhizobium sp.

C:Date: 13-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004

C:Accession: S34305

R;Relix, B.; Perret, X.; Golinowsky, W.; Pueppke, S.G.; Krishnan, H.B.; Brought

A:Description: Lipo-oligosaccharide Nod-factor signals permit rhizobial penetra

A:Reference number: S34303

A:Accession: S34305

A:Molecule type: DNA

A:Residues: 1-413 <REL>

A:Cross-references: UNIPROT:P50357; EMBL:X73362; NID:g312347; PIDN:CAAS1774.1;

C:Comment: This is one of the proteins, coded by modulation genes, that are req

C:Genetics:

A:Gene: nodC

C:Superfamily: modulation protein nodC

C:Keywords: modulation

Query Match	10.0%;	Score	199.5;	DB 1;	Length	413;			
Best Local Similarity	26.5%;	Pred. No.	3.3e-08;						
Matches	81;	Conservative	28;	Mismatches	102;	Indels	95;	Gaps	11;

QY	52	GLLAFLGGLYGAFLSAHLVAQSLFAYLHRRVAAAARGPLDAATA-RSVALTISAYQEDPAY	110
DB	8	GAVAISLYAALSTAYKGMQAIYALPTN---TTAASTPVTGSGAPPSVDVIVPCVNEDPRA	64
QY	111	LROCLASARALLYPRARLRVLMVVD---GNR-----AEDLYMVD-MPREVFADEDPATYVW	162
DB	65	LSACLASIAQDY-AGELRVYVVDGSGNRNAIIPVHDHYACDPRFRFILMPKXV-----	118
QY	163	DGNTHQPWEPAAGAGVAGAYREVEAEDPGRLAVALVTRRCVCVCAQRMGGKREVMYTA	222
DB	119	-----GKKRAEIVA	127
QY	223	FKALGDSVDYVQVCDSDTRLDPMALLELVRVLDEDPRGVAGGVDRIINPLDSDWVSFLSS	282
DB	128	IRE--SSGDLVNLVDSDTIAPDVVTKLAKM-YSVAVGAMGQLTASNRSDTLWTLFLID	184
QY	283	LRVYVAFNVNERACQSYFHCVSCISGSLGTPPGPAATQRRP-----RPSC	326
DB	185	MEYWLACNEERRAAQSRFGAVMCCC-----GPCMYRRSALLILLDKYETOLFRGRPSD	237

Qy 327 GGASRH 332
||
Db 238 FGDRH 243

RESULT 8
S12793
modulation protein nodC - Rhizobium loti
C/Species: Rhizobium loti
C/Date: 30-Sep-1993 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C/Accession: S12793
R/Collins-Emerson, J.M.; Terzaghi, E.A.; Scott, D.B.
Nucleic Acids Res. 18, 6690, 1990.
A/Title: Nucleotide sequence of Rhizobium loti nodC.
A/Reference number: S12793; MUID:91067465; PMID:2251130
A/Accession: S12793
A/Molecule type: DNA
A/Residues: 1-424 <COL>
A/Cross-references: UNIPROT:P17862; EMBL:X52958; NID:946227; PIDN:CAA37131.1; PID:g46228
C/Comment: This is one of the proteins, coded by modulation genes, that are required for
C/Genetics:
A/Gene: nodC
C/Superfamily: modulation protein nodC
C/Keywords: modulation

Query Match	8.9%	Score 176.5	DB 1	Length 424	
Best Local Similarity	26.6%	Pred. No. 2.3e-06			
Matches	83	Conservative 33	Mismatches 107	Indels 89	Gaps 13
Qy	38	WAYAAGVPLASDRYCGLLAFGLYGFLSAHLVAQSLFAYLEHRRVAAARGPLDAATARSV	97		
Db	4	PATASTVAIGCS--YALLS	:	:	:
Qy	98	ALTISAYQEDPAYLRQCLASARALLYPRARURLVMVDGNRAEDLYMVDMEFVFAEDBP	157		
Db	52	DVLIPSTNEAPRTLSDCLASTASEY-AGKLVQVYVDDGS	:	:	:
Qy	158	ATYYVDGNHYHPWEPAAAGAVGAGAYREVEADEPRLAVEALVRTRRCVCVAQRWGGKRE	217		
Db	96	LVGQVEEVGHDP-	:	:	:
Qy	218	VMTAF-KALGDSVDYVQVCDSTRLDPMALLELVRVLDEDPRVAGVGDVRIINPLDSW	276		
Db	123	AQIAAARRSCGDLVNV--DSDTILAPDVVTRTLAKM-QDOAICAAMQOLAASNRSETW	178		
Qy	277	VSFLSSRLRYVWFWNVERACQSYFHCVCISGSLGTPPGPAA	320		
Db	179	LTRLIDMEYWLACNEEERAAQARFGAVMCCC	231		

RESULT 9

JC6079 chitin synthase (EC 2.4.1.16) chsD - *Emericella nidulans*
C/Species: *Emericella nidulans*, *Aspergillus nidulans*
C/Date: 21-Jan-1997 #sequence_revision 21-Jan-1997 #text_change 20-Jun-2000
C/Accession: JC6079
R/Motoyama, T.; Fujiwara, M.; Kojima, N.; Horiuchi, H.; Okta, A.; Takagi, M.
Mol. Gen. Genet. 251, 442-450, 1996
A/Title: The *Aspergillus nidulans* genes *chsA* and *chsD* encode chitin synthases which have
A/Reference number: JC6079; MUID:96285568; PMID:8709948
A/Accession: JC6079
A/Molecule type: DNA
A/Residues: 1-1086 <MOT>
A/Cross-references: DDBJ:D83245; NID:gl688025; PIDN:BAAL1866.1; PID:gl688026
C/Comment: This enzyme functions in conidia formation.
C/Genetics:
A/Gene: *chsD*
A/Introns: 800/2

C;Superfamily: chitin synthase chs4
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 8.2%; Score 163.5; DB 2; Length 1086;
Best Local Similarity 22.3%; Pred. No. 6.9e-05;
Matches 59; Conservative 46; Mismatches 111; Indels 49; Gaps 8;

Qy	88	PLDAATARSVALTISAYQEPAYLRQCLASARALLYPRARLRVLWVDG---NRAEDLYM	144
Db	563	PYGFPLAHACL-VTCYSEGEIGRTTLDSTALDYPNSHKSIIVICDGIKKGESFST	621
Qy	145	VDMPREVFAD-----ED-----PATYVWDGNYHQWPAPAAAGAGAGAYREAEADPGLA	195
Db	622	PDIVLRMMROPIIPPEEVEAPPYVAVATGSKRHNMAKVYAGFYDYGHSIIPVMQQRVP	681
Qy	196	VEALVTRRCVCVQAQRWG-----GKR-----EVMYATFAL-----	226
Db	682	MMIIV---KCGTPAEATAAKPGNRGKRDSQIILMSFLQKMFDERMTELEYEMENGLHV	738
Qy	227	----GDSVDYVQVQCDSTRIDPMALLBELVRDLDDPRVGAVGGDVRTINPLDMSVFLSS	282
Db	739	TGIPPDYEVVLMVDATKVFPDSLTHMSANVKDPEVMGLCGETKIANKTDSWVTIQV	798
Qy	283	LRYYWAFNVERACOSYFHCVSCISG	307
Db	799	FEYFVSHHQKAFESVFGVGTCLPG	823

RESULT 10

JQ0396
 modulation protein nodC - Azorhizobium caulinodans
 C:Species: Azorhizobium caulinodans
 A:Note: host Sesbania rostrata
 C:Date: 07-Sep-1990 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
 C:Accession: JQ0396
 R:Goethals, K.; Gao, M.; Tomekpe, K.; Van Montagu, M.; Holsters, M.
 Mol. Gen. Genet. 219, 289-298, 1989
 A:Title: Common nodABC genes in Nod locus 1 of Azorhizobium caulinodans: nucle
 A:Reference number: JQ0393; MUID:90136519; PMID:2615763
 A:Accession: JQ0396
 A:Molecule type: DNA
 A:Residues: 1-395 <GO>
 A:Cross-references: UNIPROT:Q07755; GB:L18897; NID:g1293899; PIDN:AAB51164.1; P
 A:Experimental source: Strain ORS571
 C:Comment: This is one of the proteins, coded by nodulation genes, that are req
 C:Genetics:
 A:Gene: nodC
 C:Superfamily: nodulation protein nodC
 C:Keywords: nodulation

Query Match	8.0%;	Score 159.5;	DB 1;	Length 395;
Best Local Similarity	25.2%;	Pred. No. 4.6e-05;		
Matches 78;	Conservative 32;	Mismatches 101;	Indels 99;	Gaps 12;
Qy	49	DRYGLLAFGLYGLFSAHLVAQSLFAYLEHRRVAAAAARGPLDA--ATARSVALTISAYQE	106	
Db	5	DVIGLLATAAYVTLASAYKKVQ---FINVSSVTDVAGLES DALPLTPR-VDVIVPTFNE	59	
Qy	107	DPAYVROCLASARALLYPRALRLVLMVVDGNRAE-----DLYVMDMREVEADPAT	159	
Db	60	NSSTLLSECVASICADY-RGPTIVVDDGSGTNTSTFHAVCDKYASDE-RFIFVELD---	114	
Qy	160	YVMDGNTHQPWEPAAAGAVGAGAYREVEAEDPGRLA-VEALVTRTRCCVCAQRWGKKREV	218	
Db	115	-----QNKGTAAQMEAIRRT-----	129	
Qy	219	MYTAFKALGDSVDYVQVCDSTRLDPMLLELVRLVILDEDPRGVAGVGDVRLINPLDSWVS	278	
Db	130	-----DGLILNVDSDTVIDKDVTKLASSM-RAPNVGVMQGLVAKNRPERSWLT	178	
Qy	279	PLGSLRYVYAFNVVERACQSFYHCISIGSLGTFPPGPAATQRRP-----	322	
Db	179	RLIDMEYTLACNBEERTAQSRFGSVWC-----GFCAMRYRSATPLLAEYEHOTFLG	231	

```
Qy 323 RPSCGASRH 332
|||
Db 232 RPSNFGEDRH 241

RESULT 11
S61886
chitin synthase (EC 2.4.1.16) CHS4 - Neurospora crassa
N/Alternate names: chitin synthase class IV
C/Species: Neurospora crassa
C/Date: 23-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C/Accession: S61886
R/Din, A.B.; Specht, C.A.; Robbins, P.W.; Yarden, O.
Mol. Gen. Genet. 250, 214-222, 1996
A/Title: chs-4, a class IV chitin synthase gene from Neurospora crassa.
A/Reference number: S61886; MUID:96188842; PMID:8628221
A/Accession: S61886
A/Status: preliminary; nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-1195 <DIN>
A/Cross-references: UNIPROT:Q01285; EMBL:U25097; NID:g793937; PIDN:RAB03563.1; PID:g7939
C/Genetics:
A/Gene: chs-4
A/Introns: 1042/1
C/Superfamily: chitin synthase chs4
C/Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase; transmembrane

Query Match 8.0%; Score 159; DB 2; Length 1195;
Best Local Similarity 22.2%; Pred. No. 0.00018;
Matches 69; Conservative 51; Mismatches 127; Indels 64; Gaps 11;

Qy 46 LASDRYGLAFLGYG---AFLSAHLVAQLFAYLLEHRRVAAARGPLDAATARSVALTI 101
|||
Db 606 LKSDAYSSSPADGGPAGIHEAVVQPPSDWM-----PGFPLAHTICU-V 653

Qy 102 SAYQEDPAYLRQCLASARALLYPRARLVLMVVDG---NRAEDLYVDMFREVFADE--- 155
|||
Db 654 TAYSEGEVGTTLTDSIAMTDPNSHKVILVICDGIKGGHEHSTPDIILGMKDHITH 713

Qy 156 ----DPATY--VWDGNVHQHPPEAAGAGVAGAYREVEAEDPGRGLAVEALVTRRCVCA 209
|||
Db 714 PDDVEPFSYAVATGSKRHNMAKVYTFDYGTNSAIFLEKQQRVPMVMV---KCGTPA 770

Qy 210 QRW-----GGRK-----EVMYTAFAKAL-----GDSVDYVQV---C 236
|||
Db 771 EASKSKPGRNGKRSQIILMSFLQKVPFDERMTELEVENFNLKWKITGISDPFVEIVLMV 830

Qy 237 DSDTRLDPMALLEVRVLDPRVAGVGDVRLNPLDSWVSFSLSLRYWVAFNVERACQ 296
|||
Db 831 DADTKVPFDSLTHMISAMVKDPEIMGLCGETKIANKRASWYSIAIQVPEYFVSHHLAKAFE 890

Qy 297 SYFHCVSCISG 307
|||
Db 891 SVFGVGTCLPG 901

RESULT 12
S45879
chitin synthase (EC 2.4.1.16) 3 - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein YBR023c; protein YBR0305
C/Species: Saccharomyces cerevisiae
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C/Accession: S45879; S46554; S22776; A39639; S17247
R/Grivell, L.A.; de Haan, M.; Maat, M.J.; Smits, P.H.M.
submitted to the Protein Sequence Database, August 1994
A/Reference number: S45879
A/Accession: S45879
A/Molecule type: DNA
A/Residues: 1-1165 <GRI>
A/Cross-references: UNIPROT:P29465; EMBL:Z35892; NID:gs362229; PIDN:CAA84965.1; PID:gs362
A/Experimental source: strain S288C
R/Smits, P.H.M.; de Haan, M.; Maat, C.; Grivell, L.A.
```

```
Yeast 10(Suppl.A), S75-S80, 1994
A/Title: The complete sequence of a 33 kb fragment on the right arm of chromosome II fro
ly identified genes and a homologue of the SCO1 gene.
A/Reference number: S46551; MUID:94378725; PMID:8091864
A/Accession: S46554
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-1165 <SMI>
A/Cross-references: EMBL:X76078; NID:g498748; PIDN:CAA53680.1; PID:g498752
A/Experimental source: strain S288C
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
R/Bulawa, C.E.
Mol. Cell. Biol. 12, 1764-1776, 1992
A/Title: CSD2, CSD3, and CSD4, genes required for chitin synthesis in Saccharomyces cere
blum species and Xenopus laevis.
A/Reference number: S22776; MUID:92195323; PMID:1532231
A/Accession: S22776
A/Molecule type: DNA
A/Residues: 1-1162, 'L', 1164-1165 <BUL>
A/Cross-references: EMBL:M73697; NID:g172103; PIDN:AAA34844.1; PID:g172104
R/Valdivieso, M.H.; Mol, P.C.; Shaw, J.A.; Cabib, E.; Duran, A.
J. Cell Biol. 114, 101-109, 1991
A/Title: CALL, a gene required for activity of chitin synthase 3 in Saccharomyces cerevi
A/Reference number: A39639; MUID:91268144; PMID:2050737
A/Accession: A39639
A/Molecule type: DNA
A/Residues: 67-1165 <VAL>
A/Cross-references: GB:X57300; NID:g3359; PIDN:CAA40559.1; PID:g3360
C/Genetics:
A/Gene: SGD:CHS3; CALL; CSD2; MIPS:YBR023c
A/Cross-references: MIPS:YBR023c; SGD:S0000227
A/Map position: 2R
C/Function:
A/Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucosa
C/Superfamily: chitin synthase chs4
C/Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase; transmembrane
F:203-219/Domain: transmembrane #status predicted <TM1>
F:457-473/Domain: transmembrane #status predicted <TM2>
F:1018-1034/Domain: transmembrane #status predicted <TM3>
F:1035-1054/Domain: transmembrane #status predicted <TM4>
F:1060-1076/Domain: transmembrane #status predicted <TM5>
F:1084-1100/Domain: transmembrane #status predicted <TM6>

Query Match 7.9%; Score 156.5; DB 1; Length 1165;
Best Local Similarity 22.2%; Pred. No. 0.00027;
Matches 58; Conservative 48; Mismatches 88; Indels 67; Gaps 9;

Qy 101 ISAYQEDPAYLRQCLASARALLYPRARLVLMVVDG---NRAED-----LYWVDMFRE 150
|||
Db 658 VTCYSEDEEGLRTTLDLSLTDPNSHKLLMVVCDGLIKSGNDKTTPEIAGLMMDDFVT 717

Qy 151 VFADEDPATYVWDGNVHQHPPEAAGA-----VGAGAYR-----EVEAEDPGRGLAVEAL 199
|||
Db 718 PPDEVKPYSVV-----AVASGSKRHNMAKIYAGFYKYDDSTIPPENQQRVPIITI 767

Qy 200 VTRRCVCVAQRMG-----GKR--EVMYTAFA-----KAL 226
|||
Db 768 V---KCGTPAEQGAAPGNRGRKRSQIILMSFLEKITFDERMTQLEFOLLKXNIQITGLM 824

Qy 227 GDSVDYVQVCDSTRLDPMALLEVRVLDPRVAGVGDVRLNPLDSWVSFSLSLRYW 286
|||
Db 825 ADFYETVLMVDATKVPFDPALTHMVAEMVKDPLMGLCGETKIANKAQSWVTAIQVPEY 884

Qy 287 VAFNVERACQSYFHCVCISG 307
|||
Db 885 ISHHQAKAFESVFGSVTCLPG 905

RESULT 13
T42020
class IV chitin synthase - smut fungus (Ustilago maydis)
C/Species: Ustilago maydis (corn smut)
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
```

RESULT 14

T30202

C: probable chitin synthase (EC 2.4.1.16) Chs3 - Rhizopus microsporus var. oligosporus
C: Species: Rhizopus microsporus var. oligosporus
A: Variety: var. oligosporus
C: Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 09-Jul-2004
R: Accession: T30202
R: Motoyama, T.; Horiuchi, H.; Ohta, A.; Yamaguchi, I.; Takagi, M.
FEMS Microbiol. Lett. 169, 1-8, 1998
A: Title: Isolation of a class IV chitin synthase gene from a Zygomycetes fungus, Rhizopus
A: Reference number: Z20776; MUID:99068016; PMID:9851030
A: Accession: T30202
A: Status: preliminary; translated from GB/EMBL/DDBJ
A: Molecule type: DNA
A: Residues: 1-1112 <MOT>
A: Cross-references: UNIPROT:Q93787; EMBL:AB004547; NID:d1245413; PIDN:BAA3
A: Experimental source: var. oligosporus
C: Genetics:
A: Introns: 136/1; 321/2; 770/3; 817/1; 911/3
A: Note: chs3
C: Function:
A: Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucosa
A: Pathway: chitin biosynthesis
C: Superfamily: chitin synthase chs4
C: Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 7.6%; Score 151.5; DB 2; Length 1112;
Best Local Similarity 21.6%; Pred.No.0.00063;
Matches 56; Conservative 48; Mismatches 102; Indels 53; Gaps 8;

Q9 LTISAYCEDPAYLRQCASARALLYPBRLVLVMVDG-----NRADLYMVDMFREY 151
: : : | | | | | : : : : :
: : : | | | | | : : : : :

RESULT 15
A53100
hyaluronate synthase A (HasA) - Streptococcus sp. (group A)
C:Species: Streptococcus sp.
C:Date: 06-Oct-1994 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: A53100
R:Dougherty, B.A.; van de Rijn, I.
J. Biol. Chem. 269, 169-175, 1994
A:Title: Molecular characterization of hasA from an operon required for hyaluro
A:Reference number: A53100; MUID:94103204; PMID:8276791
A:Accession: A53100
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-419 <DOU>
A:Cross-references: UNIPROT:Q8NWX1
A:Note: sequence extracted from NCBI backbone (NCBIN:141683, NCBIIP:141684)
C:Superfamily: nucleation protein nodC

Search completed: March 11, 2005, 14:23:11
Job time : 30.0758 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 13:46:32 ; Search time 105.75 Seconds
(without alignments)
1375.148 Million cell updates/sec

Title: US-10-672-399-4
Perfect score: 1987
Sequence: 1 MRQDAPKPTPAARCSGLA.....TRNSPACSPSSWRPLCCVC 376

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1602	80.6	578	2 AAW26765	Aaw26765 Human hya
2	1509	75.9	582	3 AAY78135	Aay78135 Recombina
3	1509	75.9	583	2 AAW30704	Aaw30704 Mouse hya
4	1509	75.9	583	2 AAY32503	Aay32503 Hyaluro
5	1509	75.9	583	3 AAY68491	Aay68491 Mouse hya
6	1509	75.9	583	3 AAB09948	Aab09948 Murine HA
7	1509	75.9	583	6 ABP96028	Abp96028 Mouse hya
8	1509	75.9	583	7 AAE39152	Aae39152 Mouse hya
9	1422	71.6	543	2 AAW36503	Aaw36503 Human hya
10	1416	71.3	582	3 AAY78138	Aay78138 Recombina
11	1416	71.3	583	3 AAY78131	Aay78131 Recombina
12	1412.5	71.1	584	3 AAY78132	Aay78132 Recombina
13	1235	62.2	563	3 AAY78129	Aay78129 Recombina
14	993.5	50.0	577	3 AAY78128	Aay78128 Recombina
15	989	49.8	573	3 AAY78127	Aay78127 Recombina
16	873.5	44.0	552	3 AAY78133	Aay78133 Recombina
17	873.5	44.0	553	3 AAY78136	Aay78136 Recombina
18	804	40.5	552	7 ADC49213	Adc49213 Rabbit hy
19	802.5	40.4	588	7 ADD93928	Add93928 Xenopus 1
20	793	39.9	554	3 AAW50010	Aaw50010 Murine hy
21	793	39.9	554	3 AAY68493	Aay68493 Mouse hya
22	793	39.9	554	6 ABP96030	Abp96030 Mouse hya
23	793	39.9	554	6 AAE39154	Aae39154 Mouse hya
24	793	39.9	554	8 ADJ76214	Adj76214 Marker ge
25	787.5	39.6	552	3 AAY78140	Aay78140 Recombina

ALIGNMENTS

RESULT 1

AAW26765
ID AAW26765 standard; protein; 578 AA.

XX AC AAW26765;

XX 21-MAY-1998 (first entry)

XX Human hyaluronan synthase.

XX Hyaluronan synthase; HAS gene; human; hyaluronic acid; cell adhesion;
KW wound healing; vulnery; tissue repair; scar; keloid; therapy.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Binding-site 119..127
/note= "putative hyaluronan binding site"

FT Modified-site 340..343
/note= "consensus phosphorylation sequence for protein

FT Modified-site 354..356
/note= "consensus phosphorylation sequence for protein

FT Modified-site 371..374
/note= "consensus phosphorylation sequence for protein

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

FT Modified-site 378..380
/note= "consensus phosphorylation sequence for cAMP-

[illegible]

Db	242	DSDFELDPMALELVRVLDEDP	PRVCAVGGDVRILNPLDSWVSFLSLRYWVAENVERACQ	302
Qy	297	SYFHCVSCISGSLG	310	
Db	302	SYFHCVSCISGPLG	315	
RESULT 8				
AAE39152	ID	AAE39152	standard; protein; 583 AA.	
XX	AC	AAE39152;		
XX	DT	18-DEC-2003	(first entry)	
XX	XX	Mouse hyaluronan synthase (HAS) 1.		
XX	XX	Mouse; hyaluronan synthase; HA; hyaluronan synthase; HAS; dry eye syndrome;		
XX	KW	eye cell; osteoarthritis; gene therapy; enzyme.		
XX	OS	Mus sp.		
XX	XX	US2003087850-A1.		
XX	FN	08-MAY-2003.		
XX	PD	10-JUL-2001; 2001US-00902939.		
XX	PF	10-JUL-2001; 2001US-00902939.		
XX	PR	(DEHA/)	DEHAZYA P.	
XX	PA	(CHEN/)	CHEN W.	
XX	XX	Dehazya P.	Chen W;	
XX	PI	WPI; 2003-755151/71.		
XX	DR	N-PSDB; AAD59442.		
XX	XX	Dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate comprising derivatized hyaluronic acid cross linked to nucleic acid encoding hyaluronan synthase useful for treating dry eye syndrome.		
XX	PS	Claim 18; Page 15-16; 3lpp; English.		
XX	CC	The invention relates to dihydrazide derivatised hyaluronic acid (HA)/nucleic acid bioconjugate comprising derivatised HA cross linked to nucleic acid encoding hyaluronan synthase (HAS). The invention is useful for treating an eye cell of an individual. It is useful for treating dry eye syndrome and osteoarthritis of the particular joints. The invention is also useful in gene therapy. The present sequence is mouse HAS1 enzyme		
XX	SQ	Sequence 583 AA;		
Query Match				
Best Local Similarity 75.9%; Score 1509; DB 7; Length 583;				
Matches 295; Conservative 2; Mismatches 11; Indels 6; Gaps 2				
Qy	3	QQDAPKPTPAARRCSGLARRVLTTAFALLIILGLMTWAYAAGVPLASDRYGLLAFLGYL	GAF 62	
Db	2	RQDMPKPEAAARCCSGLARRALTTIFALLIILGLMTWAYAAGVPLASDRYGLLAFLGYL	GAF 61	
Qy	63	LSAHLVAQSLPAYLEHRRVAAAAR----	GPLDAATARSVALTISAYOEDPAYLRQCLAS 117	
Db	62	LSAHLVAQSLPAYLEHRRVAAAARRSLAKGPLDAATARSVALTISAYOEDPAYLRQCLTS	121	
Qy	118	ARALLYPRALRLVLMVVDGNRAEDLYMVDMEFREFADEDDPATYVWDGNVHPWEPA-AAAG	176	
Db	122	ARALLYPTHRLRLVLMVVDGNRAEDLYMVDMEFREFADEDDPATYVWDGNVHPWEPAEATG	181	
Qy	177	AVGAGAYREVEAEDPGLAVALVRLTRRCVCVAORWGKREVMYTAFKALGSDVDYVQVC	236	
Db	182	AVGEGAYREVEAEDPGLAVALVRLTRRCVCVAORWGKREVMYTAFKALGSDVDYVQVC	241	

Db 122 ABALYPHTRLRVLMVVDGNRAEDLYMVDMFREVEADEDPATYVWDGNYHQWEPFAATG 181
 QY 177 AVGAGAYREVEAEDPGRLAVALVTRTRCCVCAQWGGKREVMYTAFAKALGDSVDYVQC 236
 Db 182 AVGAGAYREVEAEDPGRLAVALVTRTRCCVCAQWGGKREVMYTAFAKALGDSVDYVQC 241
 QY 237 DSDTRLDPMALLEVRVLDDEPRVGAAGDVRILNPLDSWVSFLSLRYWVAFNVERACQ 296
 Db 242 DSDTMLDPASSVEMVKVLEEDPMVGGVGDVQILNKYDSWISFLSSVRYWMAFNTERACQ 301
 QY 297 SYFHCVCISGSLG 310
 Db 302 SYFGVCQCISGPLG 315

RESULT 11

AAV78131
 ID AAY78131 standard; protein; 583 AA.

AC AAY78131;

DT 27-APR-2000 (first entry)

DE Recombinant chimeric hyaluronate synthase modified protein #5.

KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
 KW research reagent; biochemical research; medical development; chimeric.

OS Mus sp.

OS Synthetic.

OS Chimeric.

XX JP2000004886-A.

XX 11-JAN-2000.

XX 24-JUN-1998; 98JP-00193788.

XX 24-JUN-1998; 98JP-00193788.

XX (SEBK) SEIKAGAKU KOGYO CO LTD.

XX WPI; 2000-140125/13.

XX A hyaluronate synthase modified protein - useful as a research reagent
 PT for biochemical research and medical development.

XX Claim 6; Page; 30pp; Japanese.

XX The present invention describes a recombinant protein which consists of
 CC three continuous regions (N-terminal region, internal region and C-
 CC terminal region) where one or two regions among the above three regions
 CC is selected from the three hyaluronate synthase (HAS) modified proteins
 CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
 CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
 CC is useful as a research reagent for biochemical research and medical
 CC development. The invention provides a HAS modified protein of increased
 CC or lowered activity. AAY78127 to AAY78142 represent specifically claimed
 CC examples of recombinant proteins from the present invention. N.B. The
 CC present sequence is not given in the present specification, but is
 CC derived from sequences given as specified in the claim

XX Sequence 583 AA;

XX Query Match 71.3%; Score 1416; DB 3; Length 583;

XX Best Local Similarity 87.3%; Pred. No. 4e-135;

XX Matches 274; Conservative 13; Mismatches 21; Indels 6; Gaps 2;

QY 3 QODAPKPTPAARRCCSGLARRVLTITAFALLILGLMTWYAAAGVPLASDRYGLAFLYGF 62

Db 2 RODMPKPEAARRCCSGLARRVLTITAFALLILGLMTWYAAAGVPLASDRYGLAFLYGF 61

QY 63 LSAHLVAQSLPAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQSDPAYLRQCLAS 117
 Db 62 LSAHLVAQSLPAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQSDPAYLRQCLTS 121
 QY 118 ABALYPHTRLRVLMVVDGNRAEDLYMVDMFREVEADEDPATYVWDGNYHQWEPFAATG 176
 Db 122 ABALYPHTRLRVLMVVDGNRAEDLYMVDMFREVEADEDPATYVWDGNYHQWEPFAATG 181
 QY 177 AVGAGAYREVEAEDPGRLAVALVTRTRCCVCAQWGGKREVMYTAFAKALGDSVDYVQC 236
 Db 182 AVGAGAYREVEAEDPGRLAVALVTRTRCCVCAQWGGKREVMYTAFAKALGDSVDYVQC 241
 QY 237 DSDTRLDPMALLEVRVLDDEPRVGAAGDVRILNPLDSWVSFLSLRYWVAFNVERACQ 296
 Db 242 DSDTMLDPASSVEMVKVLEEDPMVGGVGDVQILNKYDSWISFLSSVRYWMAFNTERACQ 301
 QY 297 SYFHCVCISGSLG 310
 Db 302 SYFGVCQCISGPLG 315

RESULT 12

AAV78132

ID AAY78132 standard; protein; 584 AA.

AC AAY78132;

DT 27-APR-2000 (first entry)

DE Recombinant chimeric hyaluronate synthase modified protein #6.

KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
 KW research reagent; biochemical research; medical development; chimeric.

OS Mus sp.

OS Synthetic.

OS Chimeric.

XX JP2000004886-A.

XX 11-JAN-2000.

XX 24-JUN-1998; 98JP-00193788.

XX 24-JUN-1998; 98JP-00193788.

XX (SEBK) SEIKAGAKU KOGYO CO LTD.

XX WPI; 2000-140125/13.

XX A hyaluronate synthase modified protein - useful as a research reagent
 PT for biochemical research and medical development.

XX Claim 7; Page; 30pp; Japanese.

XX The present invention describes a recombinant protein which consists of
 CC three continuous regions (N-terminal region, internal region and C-
 CC terminal region) where one or two regions among the above three regions
 CC is selected from the three hyaluronate synthase (HAS) modified proteins
 CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
 CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
 CC is useful as a research reagent for biochemical research and medical
 CC development. The invention provides a HAS modified protein of increased
 CC or lowered activity. AAY78127 to AAY78142 represent specifically claimed
 CC examples of recombinant proteins from the present invention. N.B. The
 CC present sequence is not given in the present specification, but is
 CC derived from sequences given as specified in the claim

XX Sequence 584 AA;

XX Query Match 71.1%; Score 1412.5; DB 3; Length 584;

XX Best Local Similarity 87.3%; Pred. No. 9.1e-135;

XX Matches 275; Conservative 13; Mismatches 20; Indels 7; Gaps 3;

Qy	3	QDAPKPTPAARCCSGLARRVLTIIAFALLIIGLMTWYAAGVPLASDRYGLLAFLGYGAF	62
Dd	2	RDMKPSEAAACCSGLARRALTIIFALLIIGLMTWYAAGVPLASDRYGLLAFLGYGAF	61
Qy	63	LSAHLVAQSILFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYOEDPAYLRCLAS	117
Dd	62	LSAHLVAQSILFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYOEDPAYLRCLTS	121
Qy	118	ARALLYPRARLVLMVDGNRAEDLIYMVMDFREFADEDPATYYWDGNHYQPWEPA-AAAG	176
Dd	122	ARALLYPHTRLRLVLMVDGNRAEDLIYMVMDFREFADEDPATYYWDGNHYQPWEPEATG	181
Qy	177	AVGAGAYRVEAEADPGRLAVEALVTTRRCVCVAORWGCKREVMT-AFKALGDSVDYIQV	235
Dd	182	AVGEGAYRVEAEADFCRLAVEALVTTRRCVCVAORWGCKREVMMYTAAFKALGNSVDYIQV	241
Qy	236	CUSDRTLDDPMALLELRVLVEDPPRYGAVGGDVRIINPLDMSVFLSSLIRLYWAFNVERAC	295
Dd	242	CUSDTVLDPACTIEMLRVLVEDPDQVGVGGDVIINKYDWSIFSLSVRVWFMAFNVERAC	301
Qy	296	QSYFHCVCSIGSLG	310
Dd	302	QSYFGCVOCISGPLG	316

RESULT 13
AAAY78129
ID AAAY78129 standard: protein: 563 AA.

AC AAY78129;

DT 27-APR-2000 (first entry)

DE Recombinant chimeric hyaluronate synthase modified protein #3.

Murine; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
research reagent; biochemical research; medical development; chimeric.
research reagent; biochemical research; medical development; chimeric.

OS Mu8 sp.

OS Synthetic

OS Chimeric.

PN JP2000004886-A:

11-JAN-2000.

24-JUN-1998: 98JP-00193788.

PR 24-JUN-1998: 98JP-00193788.

XX PA (SEKG) SEIKAGAKU KOGYO CO LTD.

WPI: 2000-140125/13.

AA A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.

PS Claim 4: Page: 30pp: Japanese.

The present invention describes a recombinant protein which consists of three continuous regions (N-terminal region, internal region and C-terminal region) where one or two regions among the above three regions is selected from the three hyaluronate synthase (HAS) modified proteins HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71, HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein is useful as a research reagent for biochemical research and medical development. The invention provides a HAS modified protein of increased or lowered activity. ANY78127 to ANY78142 represent specifically claimed examples of recombinant proteins from the present invention. N.B. The present sequence is not given in the present specification, but is derived from sequences given as specified in the claim

Sequence 563 AA:

Query Match 62.2%; Score 1235; DB 3; Length 563;
Best Local Similarity 83.8%; Pred. No. 1.2e-116;
Matches 249; Conservative 7; Mismatches 23; Indels 1

Qy	21	RRVUTIAFALLIGLMTWAYAAGVP-LASDRYGLLAFGLYGNFLSAHLVAQSIFAYLEHR	79
Db	11	RIIGTTLFGVSLLGITAAATVYGQFIQTQDNY-YFSFGLYGA-----LFAYLEHR	59
Qy	80	RVAAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVV	134
Db	60	RVAAAAARRSLAKGPLDAATARSVALTISAYQEDPAYLRQCLTSARALLYPHTRLRVLMVV	119
Qy	135	DGNRAEDLYMVDMPREVFADBDPATYVWDGNHYQWPEPA-AAGAVGAGAYREVEADDPGR	193
Db	120	DGNRAEDLYMVDMPREVFADBDPATYVWDGNHYQWPEPAEATGAYGEGAGAYREVEADDPGR	179
Qy	194	LAVEALVTRTRCVCVCAORWGKREVMYTAFAKALGDSVDYVQVCDSDTRLDPMLLELRV	253
Db	180	LAVEALVTRTRCVCVCAORWGKREVMYTAFAKALGDSVDYVQVCDSDTRLDPMLLELRV	239
Qy	254	LDEDPRYGAVGGDVRILNPLDSWPSFLSSRLRYWAFNVERACQSFYHCVCSIGSLG	310
Db	240	LDEDPRYGAVGGDVRILNPLDSWPSFLSSRLRYWAFNVERACQSFYHCVCSIGSLG	296

RESULT 14

AAV78128

ID AAY78128 standard; protein; 577 AA.

AC AAY78128:

DT 27-APR-2000 (first entry)

DE Recombinant chimeric hyaluronate synthase modified protein #2.

Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.

XX	05	M18 80
----	----	--------

[illegible]OS
Chimeri

PN JP2000004886-A.

11-JAN-2000

24-JUN-1998: 98JP-00193788.

PR 24-JUN-1998: 98JP-00193788.

PA (SEK) SEIKAGAKU KOGYO CO LTD.

DR WPI: 2000-140125/13.

A hyaluronate synthase modified protein - useful as a research reagent for biochemical research and medical development.

xx
PS
Claim 3: Page: 30pp: Japanese.

The present invention describes a recombinant protein which consists of three continuous regions (N-terminal region, internal region and C-terminal region) where one or two regions among the above three regions are selected from the three hyaluronate synthase (HAS) modified proteins HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71, HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 593). The protein is useful as a research reagent for biochemical research and medical development. The invention provides a HAS modified protein of increased or lowered activity. AAY78127 to AAY78142 represent specifically claimed examples of recombinant proteins from the present invention. N.B. The present sequence is not given in the present specification, but is derived from sequences given as specified in the claim.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.
OM protein - protein search, using sw model
Run on: March 11, 2005, 14:22:15 ; Search time 84.3864 Seconds
(without alignments)
1469.688 Million cell updates/sec

Title: US-10-672-399-4
Perfect score: 1987
Sequence: 1 MRQDAPKPTPAARRCSGLA.....TRWSPACSPSWRFLCCVC 376

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 1396920 seqs, 329844858 residues
Total number of hits satisfying chosen parameters: 1396920

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description	
Result No.	Score	Match	Length	ID		
1	1987	100.0	376	16	US-10-672-399-4	Sequence 4, Appli
2	1612	81.1	360	16	US-10-672-399-6	Sequence 6, Appli
3	1602	80.6	578	13	US-10-042-523-2	Sequence 2, Appli
4	1502	80.6	578	16	US-10-672-399-2	Sequence 2, Appli
5	1509	75.9	583	10	US-09-902-939-4	Sequence 4, Appli
6	1363	68.6	320	16	US-10-672-399-8	Sequence 8, Appli
7	802.5	40.4	587	13	US-10-042-523-3	Sequence 3, Appli
8	802.5	40.4	588	15	US-10-309-560-10	Sequence 10, Appli
9	793	39.9	554	10	US-09-902-939-5	Sequence 6, Appli
10	783.5	39.4	553	15	US-10-235-027-370	Sequence 370, App
11	783.5	39.4	553	15	US-10-188-832-137	Sequence 137, App
12	780.5	39.3	552	10	US-09-902-939-5	Sequence 5, Appli
13	780.5	39.3	552	14	US-10-262-526-2	Sequence 2, Appli

14	780.5	39.3	552	14	US-10-262-526-4	Sequence 4, Appli
15	308	15.5	567	9	US-09-879-959-7	Sequence 7, Appli
16	308	15.5	567	14	US-10-172-527-7	Sequence 7, Appli
17	308	15.5	567	15	US-10-309-560-12	Sequence 12, Appli
18	308	15.5	568	14	US-10-011-768B-10	Sequence 10, Appli
19	308	15.5	568	14	US-10-011-771B-10	Sequence 10, Appli
20	167.5	8.4	417	15	US-10-309-560-15	Sequence 15, Appli
21	167.5	8.4	417	15	US-10-309-560-19	Sequence 19, Appli
22	167.5	8.4	417	15	US-10-309-560-20	Sequence 20, Appli
23	167.5	8.4	417	15	US-10-309-560-21	Sequence 21, Appli
24	167.5	8.4	417	15	US-10-309-560-26	Sequence 26, Appli
25	167.5	8.4	417	15	US-10-309-560-27	Sequence 27, Appli
26	167.5	8.4	417	15	US-10-309-560-28	Sequence 28, Appli
27	167.5	8.4	417	15	US-10-309-560-29	Sequence 29, Appli
28	166.5	8.4	417	9	US-09-879-959-2	Sequence 2, Appli
29	166.5	8.4	417	14	US-10-011-768B-2	Sequence 2, Appli
30	166.5	8.4	417	14	US-10-011-771B-2	Sequence 2, Appli
31	166.5	8.4	417	14	US-10-172-527-2	Sequence 2, Appli
32	166.5	8.4	417	14	US-10-326-185-2	Sequence 2, Appli
33	166.5	8.4	417	15	US-10-309-560-2	Sequence 2, Appli
34	166.5	8.4	417	15	US-10-309-560-16	Sequence 16, Appli
35	166.5	8.4	417	15	US-10-309-560-17	Sequence 17, Appli
36	166.5	8.4	417	15	US-10-309-560-18	Sequence 18, Appli
37	166.5	8.4	417	15	US-10-309-560-22	Sequence 22, Appli
38	166.5	8.4	417	15	US-10-309-560-23	Sequence 23, Appli
39	166.5	8.4	417	15	US-10-309-560-24	Sequence 24, Appli
40	166.5	8.4	417	15	US-10-309-560-25	Sequence 25, Appli
41	159	8.0	1219	15	US-10-369-493-3212	Sequence 3212, Ap
42	156.5	7.9	1165	9	US-09-801-368-76	Sequence 76, Appli
43	156.5	7.9	1165	15	US-10-369-493-1432	Sequence 1432, Ap
44	156	7.9	1160	15	US-10-369-493-12860	Sequence 12860, A
45	150.5	7.6	750	14	US-10-156-761-12649	Sequence 12649, A

ALIGNMENTS

RESULT 1
US-10-672-399-4
; Sequence 4, Application US/10672399
; Publication No. US2005003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-4

Query Match					100.0%;	Score 1987;	DB 16;	Length 376;
Best Local Similarity					100.0%;	Pred. No. 2.4e-169;		
Matches 376;					Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MRQDAPKPTPAARRCSGLARRVITIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYL	60					
Db	1	MRQDAPKPTPAARRCSGLARRVITIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYL	60					
Qy	61	AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILASARA	120					
Db	61	AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILASARA	120					
Qy	121	LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVVDGNHYQFWPSPAAGVGA	180					
Db	121	LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVVDGNHYQFWPSPAAGVGA	180					

Qy 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db |||||
Qy 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db |||||
Qy 241 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 300
Db |||||
Qy 241 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 300
Db |||||
Qy 301 CVSCISGLTTPPGPAATQRRPRSCGASRHAGPSRTSVSGCTTRSGGTGMRGPTRRW 360
Db |||||
Qy 301 CVSCISGLTTPPGPAATQRRPRSCGASRHAGPSRTSVSGCTTRSGGTGMRGPTRRW 360
Db |||||
Qy 361 SPACSPSWRPLCCVC 376
Db |||||
Qy 361 SPACSPSWRPLCCVC 376
Db |||||

RESULT 2

US-10-672-399-6
; Sequence 6, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-6

Query Match 81.1%; Score 1612; DB 16; Length 360;
Best Local Similarity 98.7%; Pred. No. 8.7e-136;
Matches 311; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60
Db |||||
Qy 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60
Db |||||
Qy 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Db |||||
Qy 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Db |||||
Qy 121 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNYHQPWEPAAAGAVGA 180
Db |||||
Qy 121 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNYHQPWEPAAAGAVGA 180
Db |||||
Qy 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db |||||
Qy 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db |||||
Qy 241 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 300
Db |||||
Qy 241 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 300
Db |||||
Qy 301 CVSCISGLTTPPGP 315
Db |||||
Qy 301 CVSCISGLTTPPGP 315
Db |||||

RESULT 3

US-10-672-399-2
; Sequence 2, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic

Acids and Uses Thereof

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/042,523
FILING DATE: 19-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/635,552A
FILING DATE: 22-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LKS95-07
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 578 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-042-523-2

Query Match 80.6%; Score 1602; DB 13; Length 578;
Best Local Similarity 99.7%; Pred. No. 1.3e-134;
Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60
Db |||||
Qy 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60
Db |||||
Qy 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Db |||||
Qy 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Db |||||
Qy 121 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNYHQPWEPAAAGAVGA 180
Db |||||
Qy 121 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNYHQPWEPAAAGAVGA 180
Db |||||
Qy 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db |||||
Qy 181 GAYREVEAEDPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db |||||
Qy 241 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 300
Db |||||
Qy 241 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFH 300
Db |||||
Qy 301 CVSCISGL 310
Db |||||
Qy 301 CVSCISGL 310
Db |||||

RESULT 4

US-10-672-399-2
; Sequence 2, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics

```
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-2

Query Match      80.6%; Score 1602; DB 16; Length 578;
Best Local Similarity 99.7%; Pred. No. 1.3e-134;
Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFLYLG 60
DB 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFLYLG 60
QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
DB 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
QY 121 LLYPRARLRVLMVVDGNRAEDLYVMDMPREVFADEDPATYVWDGNYHQWPEAAAGVGA 180
DB 121 LLYPRARLRVLMVVDGNRAEDLYVMDMPREVFADEDPATYVWDGNYHQWPEAAAGVGA 180
QY 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMTAFKALGDSVDYVQVCDSDT 240
DB 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMTAFKALGDSVDYVQVCDSDT 240
QY 241 RLDPMALLELRVLDDEPRVGAAGVGDVRLNPLDSWVSFLSLRYWVAFNVERACQSYFH 300
DB 241 RLDPMALLELRVLDDEPRVGAAGVGDVRLNPLDSWVSFLSLRYWVAFNVERACQSYFH 300
QY 301 CVSCISGSLG 310
DB 301 CVSCISGSLG 310

RESULT 5
US-09-902-939-4
; Sequence 4, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Phillip DeHazy
; APPLICANT: William Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/OH020-USO
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 583
; TYPE: PRT
; ORGANISM: mouse
US-09-902-939-4

Query Match      75.9%; Score 1509; DB 10; Length 583;
Best Local Similarity 93.9%; Pred. No. 2.7e-126;
Matches 295; Conservative 2; Mismatches 11; Indels 6; Gaps 2;

QY 3 QODAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFLYLGAF 62
DB 2 RDMKPEARCCSGLARRALTIIIFALLILGLMTWYAAGVPLASDRYGLLAFLYLGAF 61
QY 63 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLAS 117
DB 62 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLAS 121

; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-2

Query Match      80.6%; Score 1602; DB 16; Length 578;
Best Local Similarity 99.7%; Pred. No. 1.3e-134;
Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFLYLG 60
DB 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFLYLG 60
QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
DB 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
QY 121 LLYPRARLRVLMVVDGNRAEDLYVMDMPREVFADEDPATYVWDGNYHQWPEAAAGVGA 180
DB 121 LLYPRARLRVLMVVDGNRAEDLYVMDMPREVFADEDPATYVWDGNYHQWPEAAAGVGA 180
QY 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMTAFKALGDSVDYVQVCDSDT 240
DB 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMTAFKALGDSVDYVQVCDSDT 240
QY 241 RLDPMALLELRVLDDEPRVGAAGVGDVRLNPLDSWVSFLSLRYWVAFNVERACQSYFH 300
DB 241 RLDPMALLELRVLDDEPRVGAAGVGDVRLNPLDSWVSFLSLRYWVAFNVERACQSYFH 300
QY 301 CVSCISGSLG 310
DB 301 CVSCISGSLG 310

RESULT 6
US-10-672-399-8
; Sequence 8, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 8
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-8

Query Match      68.6%; Score 1363; DB 16; Length 320;
Best Local Similarity 83.4%; Pred. No. 1.5e-113;
Matches 267; Conservative 0; Mismatches 1; Indels 52; Gaps 1;

QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
DB 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
QY 121 LLYPRARLRVLMVVDGNRAEDLYVMDMPREVFADEDPATYVWDGNYHQWPEAAAGVGA 180
DB 61 LLYPRARLRVLMVVDGNRAEDLYVMDMPREVFADEDPATYVWDGNYHQWPEAAAGVGA 120
QY 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMTAFKALGDSVDYVQVCDSDT 240
DB 121 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMTAFKALGDSVDYVQVCDSDT 180
QY 241 RLDPMALLELRVLDDEPRVGAAGVGDVRLNPLDSWVSFLSLRYWVAFNVERACQSYFH 300
DB 181 RLDPMALLELRVLDDEPRVGAAGVGDVRLNPLDSWVSFLSLRYWVAFNVERACQSYFH 240
QY 301 CVSCISGSLG 310
DB 241 CVSCISGSLG 310

RESULT 7
US-10-042-523-3
; Sequence 3, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
```


Db 176 RAVWASTFSCIMQKGGKREVMYTAFAKALGNSVDYIQVCDSDTVLDPACTIEMLRVLEB 235
Qy 257 DPRVAGVGDRILNPLDSWSFLSSRLRYWAFNVERACQSYFHCVCISGSLG 310
Db 236 DPQVGVDVQILNKYDSWISFLSSVRYWMAFNVERACQSYFCVCQISGPLG 289

RESULT 10
US-10-295-027-370
; Sequence 370, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 370
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-370

Query Match 39.4%; Score 783.5; DB 15; Length 553;
Best Local Similarity 55.1%; Pred. No. 2.5e-61;
Matches 161; Conservative 35; Mismatches 81; Indels 15; Gaps 4;
Qy 21 RRVLTAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYGAFLSAHLVAQSLFAYLEHRR 80
Db 10 RVVGTSLFALAVLGGILAAVYTGQFIHTEKHYSFGLYGAILGLHLIQLSLFAFLEHRR 69
Qy 81 VAAARG-PLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRALRLVLMVVDGNRA 139
Db 70 MRRAGQALKLPSPRRGSVALCIAAYQEDPDYLRKCLRSARISFP--DLKVMVMVVDGNRQ 127
Qy 140 EDLYMVDMFREVF-ADEDPATYVMDGNYPHQPWEPAAAGAVGAGYREVEADPGRLAYEA 198
Db 128 EDAYMLDIFHEVLGGTEQAGFFVWRNPFHEAGEGTEASLQEGMDR-----VRD 176
Qy 199 LVRTRRCVCAQRWGGKREVMYTAFAKALGNSVDYIQVCDSDTVLDPACTIEMLRVLEB 258
Db 70 MRRAGQALKLPSPRRGSVALCIAAYQEDPDYLRKCLRSARISFP--DLKVMVMVVDGNRQ 127
Qy 140 EDLYMVDMFREVF-ADEDPATYVMDGNYPHQPWEPAAAGAVGAGYREVEADPGRLAYEA 198
Db 128 EDAYMLDIFHEVLGGTEQAGFFVWRNPFHEAGEGTEASLQEGMDR-----VRD 176
Qy 199 LVRTRRCVCAQRWGGKREVMYTAFAKALGNSVDYIQVCDSDTVLDPACTIEMLRVLEB 258

Db 177 VVRSTFSCIMQKGGKREVMYTAFAKALGNSVDYIQVCDSDTVLDPACTIEMLRVLEB 236
Qy 259 RVGAVGGDRILNPLDSWSFLSSRLRYWAFNVERACQSYFHCVCISGSLG 310
Db 237 QVGVDVQILNKYDSWISFLSSVRYWMAFNVERACQSYFCVCQISGPLG 288

RESULT 11
US-10-188-832-137
; Sequence 137, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 137
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-137

Query Match 39.4%; Score 783.5; DB 15; Length 553;
Best Local Similarity 55.1%; Pred. No. 2.5e-61;
Matches 161; Conservative 35; Mismatches 81; Indels 15; Gaps 4;
Qy 21 RRVLTAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYGAFLSAHLVAQSLFAYLEHRR 80
Db 10 RVVGTSLFALAVLGGILAAVYTGQFIHTEKHYSFGLYGAILGLHLIQLSLFAFLEHRR 69
Qy 81 VAAARG-PLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRALRLVLMVVDGNRA 139
Db 70 MRRAGQALKLPSPRRGSVALCIAAYQEDPDYLRKCLRSARISFP--DLKVMVMVVDGNRQ 127
Qy 140 EDLYMVDMFREVF-ADEDPATYVMDGNYPHQPWEPAAAGAVGAGYREVEADPGRLAYEA 198
Db 128 EDAYMLDIFHEVLGGTEQAGFFVWRNPFHEAGEGTEASLQEGMDR-----VRD 176
Qy 199 LVRTRRCVCAQRWGGKREVMYTAFAKALGNSVDYIQVCDSDTVLDPACTIEMLRVLEB 258
Db 177 VVRSTFSCIMQKGGKREVMYTAFAKALGNSVDYIQVCDSDTVLDPACTIEMLRVLEB 236
Qy 259 RVGAVGGDRILNPLDSWSFLSSRLRYWAFNVERACQSYFHCVCISGSLG 310
Db 237 QVGVDVQILNKYDSWISFLSSVRYWMAFNVERACQSYFCVCQISGPLG 288

RESULT 12
US-09-902-939-5
; Sequence 5, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHaza
; APPLICANT: William Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME

```
; FILE REFERENCE: 2055/OH020-US0
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 552
; TYPE: PRT
; ORGANISM: mouse
US-09-902-939-5

Query Match      39.3%; Score 780.5; DB 10; Length 552;
Best Local Similarity 53.3%; Pred. No. 4.6e-61;
Matches 155; Conservative 47; Mismatches 72; Indels 17; Gaps 6;

Qy 21 RRVLTIAFALLILGLMTWAYAAGVP-LASDRYGLLAFLGYGAFLSAHLVAQSLFAYLEHR 79
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 80 RVAARGLPDAATARSVALTISAYQEDPAYLRQCLASARALLYPRLRLVLMVVDGNRA 139
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 70 KMKKSLETP1--KLKNTVALCIAAYQEDPDYLRKCLQSVKRLTYP--GIKVMVMDGNSD 125
Qy 140 EDLYMVDMFREVFADDPATYVWDGNYHQPWEPAAAGVAGAYREVEAEDPGRLAVEAL 199
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 126 DDLYMDIFSEVIGRDKSATYIWKNFHEK-----GPGETEESHKSSQH--VTQL 174
Qy 200 VTRRCVCVAORWGGKREVMYTAFAKALGDSVDYVQVCDSDTRLDPMALLELVRVLDEPR 259
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 175 VLSNKSICIMQKGGKREVMYTAFAALGRSDYVYQVCDSDTMDLPASSVENVKVLEEDPM 234
Qy 260 VGAVGDDVRIINPLDSSWVSLSSRLRYWAFNVERACQSYFHCVCSIGSLG 310
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 235 VGGVGGDQILNKYDWSIFLSSVRYWMAFNIERACQSYFGCVQCISGPLG 285

RESULT 13
US-10-262-526-2
; Sequence 2, Application US/10262526
; Publication No. US20030108531A1
; GENERAL INFORMATION:
; APPLICANT: Hideshige Moriya
; APPLICANT: Yuichi Wada
; TITLE OF INVENTION: GENES FOR TRANSFECTION INTO BONY TISSUES
; FILE REFERENCE: OPI418
; CURRENT APPLICATION NUMBER: US/10/262,526
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: JP 2001-367091
; PRIOR FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-526-2

Query Match      39.3%; Score 780.5; DB 14; Length 552;
Best Local Similarity 52.9%; Pred. No. 4.6e-61;
Matches 155; Conservative 48; Mismatches 73; Indels 17; Gaps 6;

Qy 19 LARVLTIAFALLILGLMTWAYAAGVP-LASDRYGLLAFLGYGAFLSAHLVAQSLFAYLE 77
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 9 ILRIIGTTLFGVSLLLGITAAYIVGYQIQTQDNY-YFSFGLYGAFSLASHLIIQSLFAFLE 67
Qy 78 HRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRLRLVLMVVDGN 137
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 68 HRKMKKSLETP1--KLKNTVALCIAAYQEDPDYLRKCLQSVKRLTYP--GIKVMVMDGNS 123
Qy 138 RAEDLYMVDMFREVFADDPATYVWDGNYHQPWEPAAAGVAGAYREVEAEDPGRLAVE 197
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 124 SEDDLYMVDIFSEVIGRDKSATYIWKNFHEK-----GPGETEESHKSSQH--VT 172
```

```
Qy 198 ALVETTRCVCVAORWGGKREVMYTAFAKALGDSVDYVQVCDSDTRLDPMALLELVRVLDE 257
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 173 QLVLSNKSICIMQKGGKREVMYTAFAALGRSDYVYQVCDSDTMDLPASSVENVKVLEED 232
Qy 258 PRVGAAGDVRILNPLDSSWVSLSSRLRYWAFNVERACQSYFHCVCSIGSLG 310
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 233 PMVGVGDDQILNKYDWSIFLSSVRYWMAFNIERACQSYFGCVQCISGPLG 285

RESULT 14
US-10-262-526-4
; Sequence 4, Application US/10262526
; Publication No. US20030108531A1
; GENERAL INFORMATION:
; APPLICANT: Hideshige Moriya
; APPLICANT: Yuichi Wada
; TITLE OF INVENTION: GENES FOR TRANSFECTION INTO BONY TISSUES
; FILE REFERENCE: OPI418
; CURRENT APPLICATION NUMBER: US/10/262,526
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: JP 2001-367091
; PRIOR FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-262-526-4

Query Match      39.3%; Score 780.5; DB 14; Length 552;
Best Local Similarity 53.3%; Pred. No. 4.6e-61;
Matches 155; Conservative 47; Mismatches 72; Indels 17; Gaps 6;

Qy 21 RRVLTIAFALLILGLMTWAYAAGVP-LASDRYGLLAFLGYGAFLSAHLVAQSLFAYLEHR 79
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 80 RVAARGLPDAATARSVALTISAYQEDPAYLRQCLASARALLYPRLRLVLMVVDGNRA 139
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 70 KMKKSLETP1--KLKNTVALCIAAYQEDPDYLRKCLQSVKRLTYP--GIKVMVMDGNSD 125
Qy 140 EDLYMVDMFREVFADDPATYVWDGNYHQPWEPAAAGVAGAYREVEAEDPGRLAVEAL 199
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 126 DDLYMDIFSEVIGRDKSATYIWKNFHEK-----GPGETEESHKSSQH--VTQL 174
Qy 200 VTRRCVCVAORWGGKREVMYTAFAKALGDSVDYVQVCDSDTRLDPMALLELVRVLDEPR 259
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 175 VLSNKSICIMQKGGKREVMYTAFAALGRSDYVYQVCDSDTMDLPASSVENVKVLEEDPM 234
Qy 260 VGAVGDDVRIINPLDSSWVSLSSRLRYWAFNVERACQSYFHCVCSIGSLG 310
Db | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Qy 235 VGGVGGDQILNKYDWSIFLSSVRYWMAFNIERACQSYFGCVQCISGPLG 285

RESULT 15
US-09-879-959-7
; Sequence 7, Application US/09879959
; Patent No. US20020160489A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H
; APPLICANT: Kumari, Kehama
; APPLICANT: Deangelis, Paul
; TITLE OF INVENTION: STREPTOCOCCUS EQUISIMILIS HYALURONAN SYNTHASE GENE AND EXPRESSION
; TITLE OF INVENTION: IN BACILLUS SUBTILIS
; FILE REFERENCE: 3554.049
; CURRENT APPLICATION NUMBER: US/09/879,959
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 09/469,200
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 10
```

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 567.
; TYPE: PRT
; ORGANISM: Chlorella virus PBCV-1
US-09-879-959-7

Query Match      15.5%; Score 308; DB 9; Length 567;
Best Local Similarity 29.3%; Pred. No. 1e-18;
Matches 89; Conservative 57; Mismatches 108; Indels 50; Gaps 10;

Qy  19 LARRVLTITAFALLIIG-----LMTWYAAAGVPLASDRYGLLAFLGFLSAHLVAQSL 72
Db  15 ITSNLIAVGGASLIILAPAITGYVLHMNIALSTINGVSAYGIFVPGFF-----LAQVL 66

Qy  73 FAYLEHRR-----VAAAARGPLDAATARSVALTISAYQEDPAYLROCLASARALLYPRARL 128
Db  67 FSELNRKRLRKWISLRPKGWNDV----RLAVIIAGYREDPYMFQKLESVRDSYGNV-A 121

Qy  129 RVLMYVDGNRAEDLYMDMFREVFADEDPATYYMDGNTHQWPMPAAAGAVGAGAYREVEA 188
Db  122 RLICVIDGDEDDMRMAAVYKAIYND-----NIKP-----EFVLCE 159

Qy  189 EDPGRLAVALVTRRCVCVQACRWGCKREVMTAFK--ALGDSVDYVQVCDSDTRLDEMA 246
Db  160 DDKEGERIDS--DFSRDICVLQPHRGKRECLYTGFLAKMDFSNNAVVLIDSDTVLEKDA 217

Qy  247 LLELVRLVDEDPVCGVGDVRIINPLDSWVSFLSRLRYVAFVNERACQSYFHCVCSCIS 306
Db  218 ILEVYPLACDPEIQAVAGECKIWN-TDTLLSLLVANRYISAFCVERSAQSFRTVQCVG 276

Qy  307 GSLG 310
Db  277 GPLG 280
```

Search completed: March 11, 2005, 14:44:37
Job time : 85.3864 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 13:48:17 ; Search time 98.2727 seconds
(without alignments)
1959.259 Million cell updates/sec

Title: US-10-672-399-4

Perfect score: 1987

Sequence: 1 MRQDAPKPTPAARRCSGLA.....TRWSPACSSWRPLCCVCV 376

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1602	80.6	578	1 HAS1_HUMAN	Q92839 homo sapien
2	1594	80.2	578	2 Q9NS49	Q9NS49 homo sapien
3	1585	79.8	577	2 O81YH3	O81YH3 homo sapien
4	1542.5	77.6	583	2 Q8S742	Q8S742 papio anubi
5	1509	75.9	583	1 HAS1_MOUSE	Q61647 mus musculus
6	1504	75.7	583	2 Q8CH93	Q8CH93 rattus norv
7	926.5	46.6	458	2 Q6T488	Q6T488 brachydanio
8	804	40.5	552	2 Q9SM29	Q9SM29 oryctolagus
9	803	40.4	557	2 Q6W9J2	Q6W9J2 xenopus lae
10	802.5	40.4	588	1 HAS1_XENLA	P13563 xenopus lae
11	793.5	39.9	553	2 Q7SR37	Q7SR37 sus scrofa
12	793	39.9	554	1 HAS3_MOUSE	Q08650 mus musculus
13	791	39.8	554	2 Q8CEB9	Q8CEB9 mus musculus
14	791	39.8	554	2 Q8CH92	Q8CH92 rattus norv
15	790.5	39.8	554	2 Q9DG40	Q9DG40 brachydanio
16	783.5	39.4	553	1 HAS3_HUMAN	Q00219 homo sapien
17	783.5	39.4	553	2 Q9GRV2	Q9GRV2 homo sapien
18	781.5	39.3	552	2 Q9SM16	Q9SM16 oryctolagus
19	780.5	39.3	552	1 HAS2_HUMAN	Q92819 homo sapien
20	780.5	39.3	552	1 HAS2_MOUSE	P70312 mus musculus
21	780.5	39.3	552	2 Q8HZJ3	Q8HZJ3 equus caball
22	779.5	39.2	552	2 Q8SQ70	Q8SQ70 sus scrofa
23	778.5	39.2	552	1 HAS2_BOVIN	Q97711 bos taurus
24	775.5	39.0	552	1 HAS2_RAT	O35776 rattus norv
25	774	39.0	552	2 Q9DG41	Q9DG41 brachydanio
26	768	38.7	552	1 HAS2_CHICK	O57424 gallus gall
27	748.5	37.7	583	1 HAS3_XENLA	O57428 xenopus lae
28	747.5	37.6	583	2 Q6AZI0	Q6AZI0 xenopus lae
29	746.5	37.6	551	1 HAS2_XENLA	O57427 xenopus lae
30	592.5	29.8	281	2 Q8WTZ0	Q8WTZ0 homo sapien
31	560.5	28.2	131	2 Q8BPN0	Q8BPN0 mus musculus

Q811Y6 rattus norv
O18792 papio anubi
Q84419 paramescium
Q9WFS9 paramescium
Q9WFT0 paramescium
Q90489 brachydanio
Q9GK14 bos taurus
Q9AQ23 brachyrihob
P04340 rhizobium 1
P04341 rhizobium m
Q6X51 sinorhizobi
P50357 rhizobium s
Q6PTX8 rhizobium s
Q9T614 rhizobium m

ALIGNMENTS

RESULT 1

HAS1_HUMAN STANDARD; PRT; 578 AA.
AC Q92839; Q14470;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE (Hyaluronic acid synthase 1) (HA synthase 1) (HuHAS1).
GN Name=HAS1; Synonyms=HAS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph node;
RX MEDLINE=96394438; PubMed=8798544; DOI=10.1074/jbc.271.38.23395;
RA Shvjan A.M., Heldin P., Butcher E.C., Yoshino T., Briskin M.J.;
RT "Functional cloning of the cDNA for a human hyaluronan-synthase";
RJ J Biol Chem. 271:23395-23399(1996)
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=96244584; PubMed=8651928; DOI=10.1006/bbrc.1996.0827;
RA Itano N., Kinata K.;
RT "Molecular cloning of human hyaluronan synthase";
RL Biochem Biophys Res Commun. 222:816-820(1996).
CC -!_FUNCTION- Plays a role in hyaluronan/hyaluronic acid (HA) synthesis. Also able to catalyze the synthesis of chito-oligosaccharide depending on the substrate.
CC -!_CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-glucuronosyl(1->3)](n) + 2n UDP.
CC -!_COFACTOR: Magnesium.
CC -!_PATHWAY: Hyaluronate synthesis.
CC -!_SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!_TISSUE SPECIFICITY: Highly expressed in ovary followed by spleen, thymus, prostate, testes and large intestine. Weakly expressed in small intestine.
CC -!_SIMILARITY: Belongs to the nodC/HAS family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL; U59269; AAC50706.1; -
CC EMBL; D84424; BAA12351.1; ALT_INIT.
CC Genew; HGNC:4818; HAS1.
CC MIM; 601463; -

```

DR GO:0005887; C: integral to plasma membrane; TAS.
DR GO:0007155; P: cell adhesion; TAS.
DR GO:0006024; P: glycosaminoglycan biosynthesis; TAS.
DR InterPro; IPR001173; Glyco_trans_2.
DR Pfam; PF00535; Glycosyltransferase; Multigene family; Transferase; Transmembrane.
KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
FT DOMAIN 1 25 Cytoplasmic (Potential).
FT TRANSMEM 26 46 1 (Potential).
FT DOMAIN 47 52 Extracellular (Potential).
FT TRANSMEM 53 73 2 (Potential).
FT DOMAIN 74 399 Cytoplasmic (Potential).
FT TRANSMEM 400 420 3 (Potential).
FT DOMAIN 421 430 Extracellular (Potential).
FT TRANSMEM 431 451 4 (Potential).
FT DOMAIN 452 457 Cytoplasmic (Potential).
FT TRANSMEM 458 478 5 (Potential).
FT DOMAIN 479 497 Extracellular (Potential).
FT TRANSMEM 498 518 6 (Potential).
FT DOMAIN 519 540 Cytoplasmic (Potential).
FT TRANSMEM 541 561 7 (Potential).
FT DOMAIN 562 578 Extracellular (Potential).
FT CONFLICT 1 2 MR -> RS (in Ref. 2).
FT CONFLICT 34 34 G -> A (in Ref. 2).
SQ SEQUENCE 578 AA; 64884 MW; 355FD54B0899E43C CRC64;

Query Match 80.6%; Score 1602; DB 1; Length 578;
Best Local Similarity 99.7%; Pred. No. 1e-116;
Matches 309; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRQDAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60
DB 1 MRQDAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60

QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
DB 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120

QY 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYWDGNYHQPWEPAAAGVGA 180
DB 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYWDGNYHQPWEPAAAGVGA 180

QY 181 GAYREVEAEDPGRLAVALVTRRCVCVAQRWGKREVMYTAFAKALGDSVDYVQVCDSDT 240
DB 181 GAYREVEAEDPGRLAVALVTRRCVCVAQRWGKREVMYTAFAKALGDSVDYVQVCDSDT 240

QY 241 RLDPMALLEVRVLDDEPRVGAAGVDVRLNPLDSWVSFLSSRLRYWYAFNVERACQSYFH 300
DB 241 RLDPMALLEVRVLDDEPRVGAAGVDVRLNPLDSWVSFLSSRLRYWYAFNVERACQSYFH 300

QY 301 CVSCISGSLG 310
DB 301 CVSCISGSLG 310

RESULT 2
Q9NS49 PRELIMINARY; PRT; 578 AA.
AC Q9NS49;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hyaluronan synthase 1.
GN Name=HAS1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E.; McCready P.M.; Skowronski E.; Viswanathan V.;
RA Burkhardt-Schultz K.; Gordon L.; Dias J.; Ramirez M.; Stillwagen S.;
RA Phan H.; Velasco N.; Do L.; Regala W.; Terry A.; Brower A.; Garnes J.;
RA Danganan L.; Erier A.; Christensen M.; Georgescu A.; Avila J.; Liu S.;

RA Andreise T., Trankheim M., Attix C., Amico-Keller G., Coefield J.,
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B.,
RA Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
RA Olsen A.S., Carrano A.V.;
RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA DOE Joint Genome Institute;
RA Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E.;
RA Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC018755; AAF87845.1; -.
DR InterPro; IPR001173; Glyco_trans_2.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR002057; Isopen N synth.
DR PROSITE; PS00237; G PROTEIN RECP F1.1; UNKNOWN 1.
DR PROSITE; PS00185; IFNS 1; UNKNOWN 1_-.
SQ SEQUENCE 578 AA; 64831 MW; 2FE3A44B0D5380FF CRC64;

Query Match 80.2%; Score 1594; DB 2; Length 578;
Best Local Similarity 99.4%; Pred. No. 4.2e-116;
Matches 308; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MRQDAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60
DB 1 MRQDAPKPTPAARCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGLY 60

QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
DB 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120

QY 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYWDGNYHQPWEPAAAGVGA 180
DB 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYWDGNYHQPWEPAAAGVGA 180

QY 181 GAYREVEAEDPGRLAVALVTRRCVCVAQRWGKREVMYTAFAKALGDSVDYVQVCDSDT 240
DB 181 GAYREVEAEDPGRLAVALVTRRCVCVAQRWGKREVMYTAFAKALGDSVDYVQVCDSDT 240

QY 241 RLDPMALLEVRVLDDEPRVGAAGVDVRLNPLDSWVSFLSSRLRYWYAFNVERACQSYFH 300
DB 241 RLDPMALLEVRVLDDEPRVGAAGVDVRLNPLDSWVSFLSSRLRYWYAFNVERACQSYFH 300

QY 301 CVSCISGSLG 310
DB 301 CVSCISGSLG 310

RESULT 3
Q81YH3 PRELIMINARY; PRT; 577 AA.
AC Q81YH3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hyaluronan synthase 1.
GN Name=HAS1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
RA Klausner R.D.; Collins F.S.; Wagner L.; Shenmen C.M.; Schuler G.D.;
RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Heish F.;
RA Diatchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.;

```

RL	Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AY463695; NAR25554.1; -.
DR	InterPro; IPR001173; Glyco trans 2.
DR	InterPro; IPR000276; GPCR Rhodpn.
DR	InterPro; IPR002057; Isoepn N synth.
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR	PROSITE; PS00105; IPNS 1; UNKNOWN_1.
SQ	SEQUENCE 583 AA; 65596 MW; EA47354C89346P94 CRC64;
Query Match	
Best Local Similarity 77.6%; Score 1542.5; DB 2; Length 583;	
Matches 299; Conservative 3; Mismatches 8; Indels 5; Gaps 1;	
QY	1 MROODAPKPTPAARRCSGLARRVLTITAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYG 60
DB	1 MTQRDTPKPTPAARRCSGLARRVLTITAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYG 60
QY	61 AFISAHILVAQSLPAYLEHRRVA-----AAARGPLDAATARSVALTISAYQEDPAYLRQCL 115
DB	61 AFISAHLLAQSLPAYLEHRRVAAAARAAARGPLDAATARSVALTISAYQEDPAYLRQCL 120
QY	116 ASARALLYPRARLURLVLMVDGNRAEDLXVDMFREVFADEDPATYYWDGNYHQPWERPAAA 175
DB	121 VSARALLYPRARLURLVLMVDGNRPEDLXVDMFREVFADEDPATYYWDGNYHQPWERPAV 180
QY	176 GAYGAGAYREVEAEDPGLRAVEALVTRTRCVCVAQRWGKGKREVMYTAFAKALGDSVDYVQV 235
DB	181 GAVGVGAYREVEAEDPGLRAVEALVTRTRCVCVAQRWGKGKREVMYTAFAKALGDSVDYVQV 240
QY	236 CDSOTRLDPMALIELVRVLDDEDPRGAVGSDVRILNPLDSWVSGFLSSLRYWVAFNVERAC 295
DB	241 CDSOTRLDPMALIELVQVLDEDPRGAVGSDVRILNPLDSWVSGFLSSLRYWVAFNVERAC 300
QY	296 QSYFHCVCISGSLG 310
DB	301 QSYFHCVCISGSLG 315
RESULT 5	
ID	HAS1_MOUSE STANDARD; PRT; 583 AA.
AC	Q61647;
DT	30-MAY-2000 (Rel. 39, Created)
DT	30-MAY-2000 (Rel. 39, Last sequence update)
DE	05-JUL-2004 (Rel. 44, Last annotation update)
DE	Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE	(Hyaluronic acid synthase 1) (HA synthase 1).
GN	Name=Has1; Synonyms=Has;
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RP	[1]
RP	SEQUENCE FROM N.A.
RP	MEDLINE=96215261; PubMed=8626618; DOI=10.1074/jbc.271.17.9875;
RA	Itano N., Kimata K.;
RA	Itano N., Kimata K.;
RT	"Expression cloning and molecular characterization of HAS protein, a
RT	eukaryotic hyaluronan synthase";
RL	J. Biol. Chem. 271:9875-9878 (1996).
RL	[2]
RP	MUTAGENESIS.
RP	MEDLINE=20085071; PubMed=10617644; DOI=10.1074/jbc.275.1.497;
RA	Yoshida M., Itano N., Yamada Y., Kimata K.;
RA	"In vitro synthesis of hyaluronan by a single protein derived from
RT	mouse HAS1 gene and characterization of amino acid residues essential
RT	for the activity.";
RL	J. Biol. Chem. 275:497-506 (2000).
CC	-I- FUNCTION: plays a role in hyaluronan/hyaluronic acid (HA)
CC	synthesis. Also able to catalyze the synthesis of chito-
CC	oligosaccharide depending on the substrate.
CC	-I- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-
CC	glucuronate = [beta-N-acetyl-D-glucosaminyl (1->4)beta-D-
CC	glucuronosyl (1->3)] (n) + 2n UDP.

OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC	Cyprinidae; Danio.
NCBI_TaxID=7955;	
RN	[1]
RP	SEQUENCE FROM N.A.
RX	PubMed=14729574;
RA	Bakkars J., Kramer C., Pothof J., Quaedvlieg N.E., Spaink H.P.,
RT	Hammerschmidt M.;
RA	"Has2 is required upstream of Rac1 to govern dorsal migration of
RT	lateral cells during zebrafish gastrulation.";
RL	Development 131:525-537(2004).
DR	ENBL; AY437407; AAR97372.1; -.
DR	ZFIN; ZDB-GENE-040218-2; has1.
DR	InterPro; IPR001173; Glyco_trans_2.
FT	NON TER 458 458
SQ	SEQUENCE 458 AA; 53188 MW; 4B421299ABBF9E6E CRC64;
Query Match	46.6%; Score 926.5; DB 2; Length 458;
Best Local Similarity	56.4%; Pred. No. 4.9e-64;
Matches 171; Conservative	50; Mismatches 75; Indels 7; Gaps 3;
QY	11 PAARRCSGLARRVLITIAFAILLIIGLMTWYAAGVPLASDRYGLLAFGLYGAFLSAHLVAQ 70
Db	5 PVUKCVSLVRAFUFTFALLVLGVMLWAYIDGFQIATSPFGIIISFGFYGVLLSLHVLIQ 64
QY	71 SLFAYLEHRRVAAAARGPDLDAARTARSVALTISAYQEDPAYLRQCCLASARALLYPRALRV 130
Db	65 SFAFVEHGMRNRK---PCSYTKTGTFISAYQEDPAYLRCLQSRYALQYSELLRI 121
QY	131 LMYVDGNRBDLYMVMFRVFADDEPATYYWDGNYHQWPAPAA---AGAGACAGAYREVE 187
Db	122 VMVVDGNSDDRYMLEFRVFADQDPCGYIWQNHYHS-WNPNGODQGAEGPDADYEVL 180
QY	188 AEDPGRLAVEALVRTRCVCVAQBWGGKREVWYTAFKALGDSDYDVYQVCDSDTLDPMAL 247
Db	181 FEDPQRLEVBEIIRTKCCVICIMQKWGGKREVWYTAFKALGSADFIVQCDSDTKLDPLAT 240
QY	248 LEDLVRLDDEPRVGAVGGDVRLNPLDSWSFSLSRLYWAFNVVERACOSYFHVCVCSIG 307
Db	241 VELCKVLESNQYKAVGGDWILNLKDSYSIFMSLSLAYWAFNIERSQQSFDCVCSIG 300
QY	308 SLG 310
Db	301 PLG 303
RESULT 8	
Q95MZ9	PRELIMINARY; PRS; 552 AA.
AC	Q95MZ9;
DT	01-DSC-2001 (TrEMBLrel. 19, Created)
DT	01-DSC-2001 (TrEMBLrel. 19, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Hyalurononic acid synthase 3.
GN	Names=HAS3;
OS	Oryctolagus cuniculus (Rabbit).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
NCBI_TaxID=9986;	
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Ohno S.;
RL	Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
DR	ENBL; AB059799; BAB3265.1; -.
DR	InterPro; IPR001173; Glyco trans 2.
SQ	SEQUENCE 552 AA; 63137 MW; 3C4A805AD5347156 CRC64;
Query Match	40.5%; Score 804; DB 2; Length 552;
Best Local Similarity	56.4%; Pred. No. 2.3e-54;
Matches 164; Conservative	34; Mismatches 79; Indels 14; Gaps 3;
QY	21 RRVLTAFALLIIGLMTWYAAGVPLASDRYGLLAFGLYGAFLSAHLVAQSLFAYLEHRR 80

Db	10	RVVGTSLFALVAGGLAAAYVTGQFTHTEKHVLSFGLYGAILGHLHLLIQSLFAFLHRR	69			
Qy	81	VAAAAAGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAE	140			
Db	70	MRRARPLKPLSPRRSVALCIAAYQEDPDYLRKLCRSAQRIATFP--DLKVVVMDGNRQE	127			
Qy	141	DLYMVDMFREV-F-ADEDPATYVMDGNYPWEPAAAGAVGAGAYREVEAEDPGRLAVAL	199			
Db	128	DAYMLDIFHEVLGTEQAGFFVWRNSNFHEAGEGETEASLQGMER-----VRV	176			
Qy	200	VRTRRCVCVAQRGGKREVMYTFAPKALGDSVDYVQVCDSDTRLDPMALLELRVLDEBPR	259			
Db	177	VRTSTFSCIMQKGGKREVMYTFAPKALGDSVDYIQVCDSDTVLDPACTIEMLRLEEDPQ	236			
Qy	260	VGAVGGDVRLNPLNDSVLSLSRLRYVWAFNVERACOSYFHCVSCISGSLG	310			
Db	237	VGGVGGDVQLNKYDSWISFLSSVRYWMAFNVERACOSYFGCVQICSGPLG	287			
RESULT 9						
Q6W9J2 PRELIMINARY; PRT; 557 AA.						
AC	Q6W9J2					
DT	05-JUL-2004	(TEMBLrel. 27, Created)				
DT	05-JUL-2004	(TEMBLrel. 27, Last sequence update)				
DT	05-JUL-2004	(TEMBLrel. 27, Last annotation update)				
DE	Hyaluronic acid synthase 3.					
OS	Xenopus laevis (African clawed frog).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;					
OC	Xenopodinae; Xenopus.					
OX	NCBI_TaxID=8155;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RX	PubMed=14667843; DOI=10.1016/S0945-053X(03)00082-9;					
RR	Vigetti D., Viola M., Gornati R., Ori M., Nardi I., Passi A.,					
RA	De Luca G., Bernardini G.;					
RT	"Molecular cloning, genomic organization and developmental expression					
RT	of the Xenopus laevis hyaluronan synthase 3.";					
RL	Matrix Biol. 24:511-517(2003).					
DR	EMBL; AY302252; AAP58398.1; -.					
DR	InterPro; IPR001173; Glyco trans. 2.					
DR	Pfam; PF00535; Glycos transf. 2; 1.					
SQ	SEQUENCE 557 AA; 61060 MW; F81478B685BF7AD5 CRC64;					
Query Match 40.4%; Score 803; DB 2; Length 557;						
Best Local Similarity 55.6%; Pred. No. 2.8e-54;						
Matches 164; Conservative 36; Mismatches 71; Indels 24; Gaps 6						
Qy	17	SGLARVLTATFALLIILGLMTWYAYAGVP-LASDRYGLLAGLYGAFLSAHLVAQSLFAY	75			
Db	7	TGL-RVLTATCLFALLVLGGFLVAVYTGQYFIHTRHH-LSGFLYGAILGLHLLSQSLP	64			
Qy	76	LEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVD	135			
Db	65	LEHRKRGGRCPSGKST--VVLCIAAYQEDPEYLKCLRSVRLSP--HLRVIMVD	119			
Qy	136	GNRAEDLYMVMFREVADBDPATYVMDGNYPWEPAAAGAVGAGAYREVEAEDPGRLA	195			
Db	120	GNTEEDRYMMDIFREVMGSEGTCCYIWDKNYHSEEGQBERG-----	163			
Qy	196	VEALVRTRRCVCVAQRGGKREVMYTFAPKALGDSVDYVQVCDSDTRLDPMALLELRVL	255			
Db	164	VQBMVRNFQVCIQKQKRGKREVTYTFALRGDSVAVYQVCDSDTVLDPACTAEMLRILE	223			
Qy	256	EDPRVAVGGDVRLNPLNDSVLSLSRLRYVWAFNVERACOSYFHCVSCISGSLG	310			
Db	224	EDPEVGGVDVQLNKYDSWISFLSSVRYWMAFNVERACOSYFGCVQICSGPLG	278			
RESULT 10						
HAS1 XENLA						

ID HAS1_XENLA STANDARD; PRT; 588 AA.
AC P13563;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE Hyaluronan synthase 1 (HA synthase 1) (XHAS1) (DG42 protein).
GN Name=HAS1;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88313363; PubMed=3410156;
RA Rosa F., Sargent T.D., Rebert M.L., Michaels G.S., Jamrich M.,
RA Grunz H., Jonas E., Winkles J.A., Dawid I.B.;
RT "Accumulation and decay of DG42 gene products follow a gradient
RT pattern during Xenopus embryogenesis.";
RL Dev. Biol. 129:114-123(1988).
CC -!- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA)
CC synthesis (By similarity). May play a role in signaling or pattern
CC formation in embryonic development.
CC -!- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-
CC glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-
CC glucuronosyl(1->3)](n) + 2n UDP.
CC -!- COFACTOR: Magnesium.
CC -!- PATHWAY: Hyaluronate synthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: Embryo.
CC -!- DEVELOPMENTAL STAGE: During gastrula and early neurula stages.
CC -!- SIMILARITY: Belongs to the nodC/HAS family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements)
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; M22249; AAA49699.1; --
DR PIR; A43740; A43740.
DR InterPro; IPR001173; Glyco_trans.2.
KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
FT DOMAIN 1 28
FT TRANSFEM 29 49
FT TRANSFEM 50 61
FT TRANSFEM 62 82
FT TRANSFEM 83 411
FT TRANSFEM 412 432
FT TRANSFEM 433 433
FT TRANSFEM 434 454
FT TRANSFEM 455 456
FT TRANSFEM 457 477
FT TRANSFEM 478 505
FT TRANSFEM 506 526
FT TRANSFEM 527 543
FT TRANSFEM 544 564
FT TRANSFEM 565 588
FT TRANSFEM 588 AA; 68522 MW; 33DA3B8E331F4CE9 CRC64;
SQ SEQUENCE 588 AA; 68522 MW; 33DA3B8E331F4CE9 CRC64;
Query Match 40.4%; Score 802.5; DB 1; Length 588;
Best Local Similarity 53.4%; Pred. No. 3.2e-54;
Matches 156; Conservative 51; Mismatches 76; Indels 9; Gaps 3;
Qy 22 RVLTIATALLILGLMTWAYAAGVPLASDRYGLAFGLYGAFLSAHLVAQSLFAYLEHRRV 81
Db RIIYISFGVLLATITAYVAEFOVLKHEAILFSLGLYGLAMLLHLMQSLFAFLFLEIRRV 87
Qy 82 AAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLVLMVVDGNRAED 141

Db 88 --NKSELPCKSFKTKTVALTIAGYQENPEYLKLECKYKVPKDKLKIILVDIGNTEDD 144
Qy 142 LYVMDPRFRVFADEDPATYVWDGNYHQPWPAPAAAGAVGAGAYREVEA---BDPRLAVEA 198
Db 145 AYMMEMFKDFHGEDVGTYYWKGNYHTVKKPE---ETNKGSCPEVSPLENDGINMVVEE 201
Qy 199 LVTRTRRCVCAQRWGGKREVMYTAFAKALGDSVDVYVQVCDSDTRLDPMALLLVLRLVLEDP 258
Db 202 LVNRKRCVCMQWGGKREVMYTAFAQIGTSVDVYVQVCDSDTKLDELTAEVEMVLESND 261
Qy 259 RVGAGVDVRLNPLDLSWVSFLSLRLRYWAFNVERACQSYFHCVSCISGSIG 310
Db 262 MYGAGGVDVRLNPLDYSFISFMSLSRLRYWAFNVERACQSYFDCVSCISGPIG 313
RESULT 11
Q75R37 PRELIMINARY; PRT; 553 AA.
AC Q75R37;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hyaluronan synthase 3.
GN Name=shas3;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Kimura N., Konno Y., Yokoo M., Sato E.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB159675; BADI8881.1; --
DR InterPro; IPR001173; Glyco_trans.2.
SQ SEQUENCE 553 AA; 63110 MW; C97EBE911723B44C CRC64;
Query Match 39.9%; Score 793.5; DB 2; Length 553;
Best Local Similarity 55.8%; Pred. No. 1.5e-53;
Matches 163; Conservative 34; Mismatches 80; Indels 15; Gaps 4;
Qy 21 RRVLTATALLILGLMTWAYAAGVPLASDRYGLAFGLYGAFLSAHLVAQSLFAYLEHRR 80
Db 10 RVVGTSLFALAVLGGILAAVTVGTQFTHTEKHVLSFGLYGAILGLHLLIQSLPAFLHRR 69
Qy 81 VAAAARG-PLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLVLMVVDGNRA 139
Db 70 MRRAGRPLKLPSPQRSVALCIAAYQEDPDYLRKLSAQRIAPP--DLKVMVVDGNRQ 127
Qy 140 EDLYVMDPREVF-ADEDPATYVWDGNYHQPWPAPAAAGAVGAGAYREVEAEDPRLAVEA 198
Db 128 EDAYMLDIFHEVLGGNEQAGFFVWRSNFHEAGGETEASLQEGMDR-----VRN 176
Qy 199 LVTRTRRCVCAQRWGGKREVMYTAFAKALGDSVDVYVQVCDSDTRLDPMALLLVLRLVLEDP 258
Db 177 VVRATFSCIMQWGGKREVMYTAFAKALGDSVDVYVQVCDSDTVLDPACTFEMLRVLEDP 236
Qy 259 RVGAGVDVRLNPLDLSWVSFLSLRLRYWAFNVERACQSYFHCVSCISGSIG 310
Db 237 QVGVGVDVQILNKYDSWISFLSSVRYWAFNVERACQSYFGCVQCISGPIG 288
RESULT 12
ID HAS3_MOUSE STANDARD; PRT; 554 AA.
AC O08650;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hyaluronan synthase 3 (EC 2.4.1.212) (Hyaluronate synthase 3)
DE Hyaluronan synthase 3 (HA synthase 3).
GN Name=Has3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

236 DPOVGVGDGVDIILNKYDWSIFLSSVRYWAFNVERACQSYFGCVQCISGPIG 20

RESULT 13

Q8CEB9 PRELIMINARY; PRT; 554 AA.

ID Q8CEB9

AC Q8CEB9

CD Q8CEB9

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched library, clone:4732404L04 product:similar to DG42III.

GN Name=Ha3;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OC NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Skin;

RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-4

RX Carninci P., Hayashizaki Y.;

RT "High-efficiency full-length cDNA cloning.";

RT Meth. Enzymol. 303:19-44(1999).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Skin;

RC MEDLINE=21085560; PubMed=11217851; DOI=10.1038/35055500;

RX RIKEN FANTOM Consortium;

RT "Functional annotation of a full-length mouse cDNA collection.";

RT Nature 409:685-690(2001).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Skin;

RC The FANTOM Consortium;

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";

RT Nature 420:563-573(2002).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Skin;

RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;

RX Carninci P., Shibata K., Hayatsu N., Sugahara Y., Shibata K., Itoh K., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes"

RT Genome Res. 10:1617-1630(2000).

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Skin;

RC MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;

RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hagiwara M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;

RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multiplex capillary sequencer.";

RT Genome Res. 10:1757-1771(2000).

RN [6]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Skin;

RC Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Hayashida K., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T., Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Numazaki R., Ohno M., Ohashi N., Okazaki Y., Saito R., Saifoh H., Sakai C., Sakai K., Sakazume N., Sano H.,

Result No.	Score	Query #		Length	DB	ID	Description
		Match					
1	1071	100.0		1071	18	US-10-672-399-3	Sequence 3, Appli
2	953.4	89.0		1083	18	US-10-672-399-5	Sequence 5, Appli
3	927.4	86.6		1065	18	US-10-672-399-7	Sequence 7, Appli
4	927.4	86.6		1737	18	US-10-672-399-1	Sequence 1, Appli
5	927.4	86.6		2116	13	US-10-042-523-1	Sequence 1, Appli
6	678.2	63.3		1752	10	US-09-902-939-1	Sequence 1, Appli
7	422	39.4		662	18	US-10-363-345A-13975	Sequence 13975, A
8	422	39.4		662	18	US-10-363-345A-13976	Sequence 13976, A
C 8	393.8	36.8		662	18	US-10-363-345A-13973	Sequence 13973, A
C 9	393.8	36.8		662	18	US-10-363-345A-13974	Sequence 13974, A
10	338.6	31.6		490	10	US-09-918-595-27210	Sequence 27210, A
11	338.6	31.6		490	10	US-09-918-595-27210	Sequence 27210, A

Db 121 GCCCGCGGGTGGCCCTCCGATCGCTACGGCCCTCTGGCCCTCTACGGG 180
Qy 181 GCCTTCCTTTAGCGCACTGGTGGCGAGAGCCTCTTTCGGGTACTTGGAGACCGGGCG 240
Db 181 GCCTTCCTTTAGCGCACTGGTGGCGAGAGCCTCTTTCGGGTACTTGGAGACCGGGCG 240
Qy 241 GTGGCGGGCGGGCGGGCGGGCGGGTGGATGACGACACCGCGCGAGTGGCGCTGACC 300
Db 241 GTGGCGGGCGGGCGGGCGGGCGGGTGGATGACGACACCGCGCGAGTGGCGCTGACC 300
Qy 301 ATCTCCGGCTTACCAGAGAGACCCCGCTACCTGGCCAGTGCCTGGCGCTCCGCCCGGCC 360
Db 301 ATCTCCGGCTTACCAGAGAGACCCCGCTACCTGGCCAGTGCCTGGCGCTCCGCCCGGCC 360
Qy 361 CTGCTGTATACCGCGCGCGCTGGCTGCTCTCATGTGGTGGATGGCAACCGCGCCGAG 420
Db 361 CTGCTGTATACCGCGCGCGCTGGCTGCTCTCATGTGGTGGATGGCAACCGCGCCGAG 420
Qy 421 GACCTCTACATGCTGCAATGTTCCGCGAGGTCTTCCGTGACGAGGACCCCGCCACGTAC 480
Db 421 GACCTCTACATGCTGCAATGTTCCGCGAGGTCTTCCGTGACGAGGACCCCGCCACGTAC 480
Qy 481 GTGTGGGACGCAACTACACAGCCCTGGGAAACCGCGCGCGGGCGCGGTGGGGGCC 540
Db 481 GTGTGGGACGCAACTACACAGCCCTGGGAAACCGCGCGCGGGCGCGGTGGGGGCC 540
Qy 541 GGAGCCTATCCGGAGGTGGAGCGGAGGATCTTGGCGGGCTGGCAGTGGAGGCGCTGGTG 600
Db 541 GGAGCCTATCCGGAGGTGGAGCGGAGGATCTTGGCGGGCTGGCAGTGGAGGCGCTGGTG 600
Qy 601 AGGACTCCAGGTGGGTGGCGCGAGCGCTGGGGCGCGAGCGGAGTCACTGTAC 660
Db 601 AGGACTCCAGGTGGGTGGCGCGAGCGCTGGGGCGCGAGCGGAGTCACTGTAC 660
Qy 661 ACAGCCTTCAAGGCGCTCGAGATTCGGTGACTACGTGAGGTCTGTGACTCCGACACA 720
Db 661 ACAGCCTTCAAGGCGCTCGAGATTCGGTGACTACGTGAGGTCTGTGACTCCGACACA 720
Qy 721 AGGTGGACCCCATGGCACTGCTGGAGCTCTGGCGGGTACTTGGACGAGACCCCGGGTA 780
Db 721 AGGTGGACCCCATGGCACTGCTGGAGCTCTGGCGGGTACTTGGACGAGACCCCGGGTA 780
Qy 781 GGGGCTGTGTGGGAGCGTGGATCTTAACCTCTGGACTCTGGGTGAGTCTTCTTA 840
Db 781 GGGGCTGTGTGGGAGCGTGGATCTTAACCTCTGGACTCTGGGTGAGTCTTCTTA 840
Qy 841 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTACAGACTACTTCCAC 900
Db 841 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTACAGACTACTTCCAC 900

RESULT 2

US-10-672-399-5
; Sequence 5, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A994635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25

; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-5

Query Match 89.0%; Score 953.4; DB 18; Length 1083;
Best Local Similarity 94.5%; Pred. No. 1.9e-227;
Matches 1023; Conservative 0; Mismatches 1; Indels 59; Gaps 1;

Qy 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGACGCCCGCCGCTGCTCGGGCTTGACC 60
Db 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGACGCCCGCCGCTGCTCGGGCTTGACC 60
Qy 61 CGGAGGGTGTGACCATGCGCTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
Db 61 CGGAGGGTGTGACCATGCGCTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
Qy 121 GCCGCCGGGGTGCCTGGCTTCGATCGCTACCGGCTCTCTGGGCTTCGGGCTTACGGG 180
Db 121 GCCGCCGGGGTGCCTGGCTTCGATCGCTACCGGCTCTCTGGGCTTCGGGCTTACGGG 180
Qy 181 GCCTTCCTTTACGGCACTTGGTGGCGCAGAGCCTCTTTCGGGTACTTGGAGCAACGGCGG 240
Db 181 GCCTTCCTTTACGGCACTTGGTGGCGCAGAGCCTCTTTCGGGTACTTGGAGCAACGGCGG 240
Qy 241 GTGGCGGGCGGGCGGGGGCGGCTGGATGACGACCCCGCGCGAGTGGCGCTGACC 300
Db 241 GTGGCGGGCGGGCGGGGGCGGCTGGATGACGACCCCGCGCGAGTGGCGCTGACC 300
Qy 301 ATCTCCCGCTTACAGAGAGGACCCCGCTACTCTGGCGCAGTGCCTGGGCTTCGGCGGCC 360
Db 301 ATCTCCCGCTTACAGAGAGGACCCCGCTACTCTGGCGCAGTGCCTGGGCTTCGGCGGCC 360
Qy 361 CTGCTGTACCCCGCGCGCGCTGGCTGCTCTCATGTGGTGGATGGCAACCGCGCCGAG 420
Db 361 CTGCTGTACCCCGCGCGCGCTGGCTGCTCTCATGTGGTGGATGGCAACCGCGCCGAG 420
Qy 421 GACCTCTACATGTTCCGACATGTTCCGCGAGGTCTTCCGTGACGAGGACCCCGCCACGTAC 480
Db 421 GACCTCTACATGTTCCGACATGTTCTCCGCGAGGTCTTCCGTGACGAGGACCCCGCCACGTAC 480
Qy 481 GTGTGGGACGCAACTACACAGCCCTGGGAAACCGCGCGCGGGCGGGCGGTGGGGGCC 540
Db 481 GTGTGGGACGCAACTACACAGCCCTGGGAAACCGCGCGCGGGCGGGCGGTGGGGGCC 540
Qy 541 GGAGCCTATCCGGAGGTGGAGCGGAGGATCTTGGCGGGTGGCAGTGGAGGCGCTGGTG 600
Db 541 GGAGCCTATCCGGAGGTGGAGCGGAGGATCTTGGCGGGTGGCAGTGGAGGCGCTGGTG 600
Qy 601 AGGACTCCAGGTGGGTGGCGCGAGCGCTGGGGCGCGAGCGGAGTCACTGTAC 660
Db 601 AGGACTCCAGGTGGGTGGCGCGAGCGCTGGGGCGCGAGCGGAGTCACTGTAC 660
Qy 661 ACAGCCTTCAAGGCGCTCGGAGATTCGGTGGATTCGCTGGAGTCTGTGACTCCGACACA 720
Db 661 ACAGCCTTCAAGGCGCTCGGAGATTCGGTGGATTCGCTGGAGTCTGTGACTCCGACACA 720
Qy 721 AGGTGGACCCCATGGCACTGCTGGAGCTCTGGGGGTACTTGGACGAGGACCCCGGGTA 780
Db 721 AGGTGGACCCCATGGCACTGCTGGAGCTCTGGGGGTACTTGGACGAGGACCCCGGGTA 780
Qy 781 GGGGCTGTGTGGGAGCGTGGATCTTAACCTCTGGACTCTGGGTGAGTCTTCTTA 840
Db 781 GGGGCTGTGTGGGAGCGTGGATCTTAACCTCTGGACTCTGGGTGAGTCTTCTTA 840
Qy 841 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTACAGACTACTTCCAC 900
Db 841 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTACAGACTACTTCCAC 900

```
QY 901 TGTGTATCTCTGATCAGCGGTTCTCT----- 926
Db 901 TGTGTATCTCTGATCAGCGGTTCTCTAGAACTCTGCCAGGCCCCCAGGAGCAGCGATG 960
QY 927 -----AGGTACACCTCCAGGTCCTCTCTACTCAGAC 961
Db 961 ATGCCCTCAATCTCTGCCCGCCGTGAGGTACACCTCCAGGTCCTCTACTCAGAC 1020
QY 962 GCCCTCGTCTCTCTCTGCGGTGCTGAGCCAGCAGACAGCTGGTCCAAAGTCTCTACTTCG 1021
Db 1021 GCCCTCGTCTCTCTGCGGTGCTGAGCCAGCAGACAGCTGGTCCAAAGTCTCTACTTCG 1080
QY 1022 TGA 1024
Db 1081 TGA 1083
```

RESULT 3

```
US-10-672-399-7
; Sequence 7, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 7
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-7
```

```
Query Match 86.6%; Score 927.4; DB 18; Length 1065;
Best Local Similarity 99.9%; Pred. No. 5.4e-221;
Matches 928; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGAGACAGCAGGAGCGGCCCAAGCCCACTCTCTGAGCGCGCGCTGTCTCGGCTTGCC 60
Db 1 ATGAGACAGCAGGAGCGGCCCAAGCCCACTCTCTGAGCGCGCGCTGTCTCGGCTTGCC 60
QY 61 CGGAGGGTGCTGACCATCGCCTTCCGCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
Db 61 CGGAGGGTGCTGACCATCGCCTTCCGCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
QY 121 GCCCGGGGGTGCGCGTCCGATCGCTACGCGCTCTGGGCTCTGGGCTCTACCGG 180
Db 121 GCCCGGGGGTGCGCGTCCGATCGCTACGCGCTCTGGGCTCTGGGCTCTACCGG 180
QY 181 GCCTTCTTTTACGCGCAGCTGTGGTGGCGAGAGCTCTTTCGGGTACCTGGAGACCGCGGG 240
Db 181 GCCTTCTTTTACGCGCAGCTGTGGTGGCGAGAGCTCTTTCGGGTACCTGGAGACCGCGGG 240
QY 241 GTGGCGGGCGGGCGCGGGGCGCTGGATGCGACGCGCGCGAGTGCTGGCGCTGACC 300
Db 241 GTGGCGGGCGGGCGCGGGGCGCTGGATGCGACGCGCGCGAGTGCTGGCGCTGACC 300
QY 301 ATCTCCGCTTACACAGAGAGACCCCGGTACCTTGCGCCAGTGTCTGGCGTCCGCGCGCC 360
Db 301 ATCTCCGCTTACACAGAGAGACCCCGGTACCTTGCGCCAGTGTCTGGCGTCCGCGCGCC 360
QY 361 CTGCTGTATACCGCGCGCGGTGCGGTCTCTATGTTGGTGGATGGCAACCGCGCCGAG 420
Db 361 CTGCTGTATACCGCGCGCGGTGCGGTCTCTATGTTGGTGGATGGCAACCGCGCCGAG 420
QY 421 GACCTCTACATGGTTCGATGTTCCGCGAGGTCTTTCGGTACGAGGACCCCGCCACGTAC 480
Db 421 GACCTCTACATGGTTCGATGTTCCGCGAGGTCTTTCGGTACGAGGACCCCGCCACGTAC 480
```

RESULT 4

```
US-10-672-399-1
; Sequence 1, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-1
```

```
Query Match 86.6%; Score 927.4; DB 18; Length 1737;
Best Local Similarity 99.9%; Pred. No. 5.4e-221;
Matches 928; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 ATGAGACAGCAGGAGCGGCCCAAGCCCACTCTCTGAGCGCGCGCTGTCTCGGCTTGCC 60
Db 1 ATGAGACAGCAGGAGCGGCCCAAGCCCACTCTCTGAGCGCGCGCTGTCTCGGCTTGCC 60
QY 61 CGGAGGGTGCTGACCATCGCCTTTCGCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
Db 61 CGGAGGGTGCTGACCATCGCCTTTCGCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
QY 121 GCCCGCGGGGTGCGCGTCCGATCGCTTACGCGCTCTCTGGGCTCTCGGCTCTACCGG 180
Db 121 GCCCGCGGGGTGCGCGTCCGATCGCTTACGCGCTCTCTGGGCTCTCGGCTCTACCGG 180
QY 181 GCCTTCTTTTACGCGCAGCTGTGGTGGCGAGAGCTCTTTCGGGTACCTGGAGACCGCGGG 240
Db 181 GCCTTCTTTTACGCGCAGCTGTGGTGGCGAGAGCTCTTTCGGGTACCTGGAGACCGCGGG 240
```

```
Qy 241 GTGGCGGCGGCGCGCGGCGCGCTGGATGAGCCACCGCGCAGTGTGGCGCTGACC 300
Db 241 GTGGCGGCGGCGCGCGGCGCGCTGGATGAGCCACCGCGCAGTGTGGCGCTGACC 300
Qy 301 ATCTCCGCTTACACGAGGAGCCCGCGTACCTGGCCAGTGCCTGGCGTCCGCGCGCGCC 360
Db 301 ATCTCCGCTTACACGAGGAGCCCGCGTACCTGGCCAGTGCCTGGCGTCCGCGCGCGCC 360
Qy 361 CTGCTGTATACCGCGCGCGCTGGCGTCTCTCATGTGTGGATGGCAACCGCGCGGAG 420
Db 361 CTGCTGTATACCGCGCGCGCTGGCGTCTCTCATGTGTGGATGGCAACCGCGCGGAG 420
Qy 421 GACCTCTACATGTGTGATGTCGCGAGGTCTTCGTGACGAGGACCCCGCCACGTAC 480
Db 421 GACCTCTACATGTGTGATGTCGCGAGGTCTTCGTGACGAGGACCCCGCCACGTAC 480
Qy 481 GTGTGGGACGGCAACTACACAGCCCTGGGAAACCGCGCGCGCGCGCGGTGGGGGCC 540
Db 481 GTGTGGGACGGCAACTACACAGCCCTGGGAAACCGCGCGCGCGCGCGGTGGGGGCC 540
Qy 541 GGAGCTATCGGAGGTGGAGCGGAGATCTTGGCGCGCTGGCAGTGGAGGCGCTGGTG 600
Db 541 GGAGCTATCGGAGGTGGAGCGGAGATCTTGGCGCGCTGGCAGTGGAGGCGCTGGTG 600
Qy 601 AGGACTCGCAGGTGGTGTGGCGCAGCGCTGGGGCGGCGCGCGCGCGAGTTCATGTAC 660
Db 601 AGGACTCGCAGGTGGTGTGGCGCAGCGCTGGGGCGGCGCGCGCGCGAGTTCATGTAC 660
Qy 661 ACAGCCTTCAAGCGGCTCGGAGATTCGGTGACCTACGTGCAAGTCTGTGACTCGGACACA 720
Db 661 ACAGCCTTCAAGCGGCTCGGAGATTCGGTGACCTACGTGCAAGTCTGTGACTCGGACACA 720
Qy 721 AGGTGGACCCCATGGCACTGCTGAGAGTCTGTGGGGTACTGGACGAGGACCCCGGGTA 780
Db 721 AGGTGGACCCCATGGCACTGCTGAGAGTCTGTGGGGTACTGGACGAGGACCCCGGGTA 780
Qy 781 GGGGCTGTGTGGGAGTGGGATCCTTAACCTCTGGAATCTGGGTCAGCTTCTCTTA 840
Db 781 GGGGCTGTGTGGGAGTGGGATCCTTAACCTCTGGAATCTGGGTCAGCTTCTCTTA 840
Qy 841 AGCAGCTGCGATACTGGGTAGCTTCAATGTGGAGCGGCTGTGTGAGCTACTTCCAC 900
Db 841 AGCAGCTGCGATACTGGGTAGCTTCAATGTGGAGCGGCTGTGTGAGCTACTTCCAC 900
Qy 901 TGTGTATCTGATCAGCGGTTCTTAGG 929
Db 901 TGTGTATCTGATCAGCGGTTCTTAGG 929
```

RESULT 5

```
US-10-042-523-1
; Sequence 1, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J. Hyaluronan Syntheses, Nucleic
; TITLE OF INVENTION: Mammalian Hyaluronan Syntheses, Nucleic
; ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/042,523
```

```
; FILING DATE: 19-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1769
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-042-523-1
```

```
Query Match      86.6%; Score 927.4; DB 13; Length 2116;
Best Local Similarity 99.9%; Pred. No. 5.3e-221; Indels 0; Gaps 0;
Matches 928; Conservative 0; Mismatches 1;

Qy 1 ATGAGACAGCAGGACGCGCCAAAGCCACCTCTGACGCGCGCGCTCTCGGCGCTGGCC 60
Db 36 ATGAGACAGCAGGACGCGCCAAAGCCACCTCTGACGCGCGCGCTCTCGGCGCTGGCC 95
Qy 61 CGAGGGTGTGACCATGCGCTTTCGCGCTGCTCATCTGCGGCTCATGACTGGGCTTAC 120
Db 96 CGAGGGTGTGACCATGCGCTTTCGCGCTGCTCATCTGCGGCTCATGACTGGGCTTAC 155
Qy 121 GCGCGCGGGTGGCGCTGCGCTCCGATCGCTAGCGGCTCTGCGCTTTCGCGCTTACGGG 180
Db 156 GCGCGCGGGTGGCGCTGCGCTCCGATCGCTAGCGGCTCTGCGCTTTCGCGCTTACGGG 215
Qy 181 GCCTTCCTTTTCAGCGCACCTGTGGCGCAGAGCCTTTCGCGTACCTGGAGCACCGGCGG 240
Db 216 GCCTTCCTTTTCAGCGCACCTGTGGCGCAGAGCCTTTCGCGTACCTGGAGCACCGGCGG 275
Qy 241 GTGGCGGCGCGCGCGCGGGCGCTGGATGAGACCAACCGCGCGCAGTGTGGCGCTGACC 300
Db 276 GTGGCGGCGCGCGCGGGCGCGCTGGATGAGACCAACCGCGCGCAGTGTGGCGCTGACC 335
Qy 301 ATCTCCGCTTACACGAGGAGCCCGCGTACCTGGCCAGTGCCTGGCGCTCCGCGCGCGCC 360
Db 336 ATCTCCGCTTACACGAGGAGCCCGCGTACCTGGCCAGTGCCTGGCGCTCCGCGCGCGCC 395
Qy 361 CTGCTGTACCCCGCGCGCGCGCTGCGCTCTCATGTGTGGATGGCAACCGCGCGGAG 420
Db 396 CTGCTGTACCCCGCGCGCGCGCTGCGCTCTCATGTGTGGATGGCAACCGCGCGGAG 455
Qy 421 GACCTCTACATGTGTGACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCCACGTAC 480
Db 456 GACCTCTACATGTGTGACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCCACGTAC 515
Qy 481 GTGTGGGACGGCAACTACACAGCCCTGGGAAACCGCGCGCGCGCGCGGTGGGGGCC 540
Db 516 GTGTGGGACGGCAACTACACAGCCCTGGGAAACCGCGCGCGCGCGCGGTGGGGGCC 575
Qy 541 GGAGCTTATCGGAGGTGGAGCGGAGGATCTTGGCGCGCTGGCAGTGGAGGCGCTGGTG 600
Db 576 GGAGCTTATCGGAGGTGGAGCGGAGGATCTTGGCGCGCTGGCAGTGGAGGCGCTGGTG 635
Qy 601 AGGACTCGCAGGTGGTGTGGCGCAGCGCTGGGGCGGCAACCGCGCGAGTTCATGTAC 660
Db 636 AGGACTCGCAGGTGGTGTGGCGCAGCGCTGGGGCGGCAACCGCGCGAGTTCATGTAC 695
```



```
QY 661 ACAGCCTTCAAGCGCTCGAGATTCTGGTGGACTAGTGGAGGTCTGTGACTCGGACACA 720
Db 696 ACAGCCTTCAAGCGCTCGAGATTCTGGTGGACTAGTGGAGGTCTGTGACTCGGACACA 755
QY 721 AGGTTGGACCCCATGGCACTGTGGAGCTCGTGGGGTACTTGGACGAGGACCCCGGGTA 780
Db 756 AGGTTGGACCCCATGGCACTGTGGAGCTCGTGGGGTACTTGGACGAGGACCCCGGGTA 815
QY 781 GGGGCTGTTGGTGGGACGTGGGATCTTAACCTCTTGGACTCTTGGGTCAAGCTTTCCTA 840
Db 816 GGGGCTGTTGGTGGGACGTGGGATCTTAACCTCTTGGACTCTTGGGTCAAGCTTTCCTA 875
QY 841 ACAGCCTCGGATCTGGGTAGGCTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCCAC 900
Db 876 ACAGCCTCGGATCTGGGTAGGCTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCCAC 935
QY 901 TGTGTATCTCGATCAGCGGTTCTTAGG 929
Db 936 TGTGTATCTCGATCAGCGGTTCTTAGG 964

RESULT 6
US-09-902-939-1
; Sequence 1, Application US/0902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHazy
; APPLICANT: William Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/OH020-USO
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1752
; TYPE: DNA
; ORGANISM: mouse
US-09-902-939-1

Query Match 63.3%; Score 678.2; DB 10; Length 1752;
Best Local Similarity 84.5%; Pred. No. 4,6e-159; Indels 18; Gaps 2;
Matches 793; Conservative 0; Mismatches 128;

QY 10 CAGGACGGCCCAAGCCCACTCTCGAGCCCGCGCTGCTCGGCTTGGCCCGGAGGGTG 69
Db 7 CAGGACATGTCACCAAGCCCTCAGAGGAGCGCTTGTCTCTGCTTGGCTTGGCCAGGCGACA 66
QY 70 CTGACCATGCGCTTTCGCCCTGCTCATCTCTGGGCTCATGACTGGGCTTACGCCCGGG 129
Db 67 CTCAGATCATCTTTGGCCCTGCTCATCTCTGGGCTCATGACTGGGCTTACGCCCGAGGC 126
QY 130 GTGCGCTGGCTTCGATCGCTACGGCTTCCTGGCTTGGCTTACGGGCTTCTT 189
Db 127 GTTCTCTGCTTTCAGATCGCTATGACTCTCTGGCTTGGCTTATGGGCAATCTCTC 186
QY 190 TCAGCGCACCTGCTGGCGCAGAGCTCTTCGGGTACCTTGGAGCACCGCGGGTGGCGG 249
Db 187 AGCGCACACCTAGTGGCACAGAGCTCTTCGCTTACCTTGGAGCACCGAGGTTGACGG 246
QY 250 GCGGCGCG-----GGGCGGCTGGATGACGACACCGCGCGCATGTGGCG 294
Db 247 GCTGCGCGCGCTCTCTTGGCGAAGGGGCCCTGGATGGCGGCTACCTGACGAGCGTGGCA 306
QY 295 CTGACCATCTCGCTTACAGAGAGACCCCGGTACTCTGCGCATGCTTGGCGTTCGCG 354
Db 307 CTCACCATCTCAGCTTACCAAGAGATCCCGCTTACCTTGGCGCATGCTTACCTTCGCG 366
QY 355 CGCGCCTCTGTACCGCGCGCGGCTGCGCTCTCATGTGTGGTGGTGAACCAACGC 414
Db 367 CGCGCCTTGTGTATACCGCACACAGAGTTACCGTGTCTCATGTGTGGTGAACCAACGC 426
QY 415 GCCGAGGACCTCTACATGCTCGACATGTTTCGCGAGGTTCTTCGCTGACGAGGACCCCGC 474
```

```
Db 427 GCTGAGGATCTGTATCATGTGGACATGTTCCGAGAAGTCTTCCCGATGAGGACCCCGCC 486
QY 475 AGCTACGTGTGGAGCGGCAACTACACAGCCCTGGGAACCCCGCGCGG--CGGCGCG 531
Db 487 ACTTATGTGTGGATGGCACTACCATCAGCCCTGGGAACCCAGCGGAGGCTACGGGCGT 546
QY 532 GTGGCGCGCGGACCTATCGGAGGTGGAGCGGAGGATCTCTGGGCGGCTGGCAGTGGAG 591
Db 547 GTCCGTGAAGGTGCCCTACCGGAGGTGGAGCGGAGGACCCCGCGGCTTGGCGGTGGAG 606
QY 592 GGGCTGGTGAAGACTTCGACAGGTGCGTGTGCGTGGCGCAGCGCTGGGCGGCGAAGCCGAG 651
Db 607 GCGCTGGTGAAGAACGACAGGTGCGTGTGCGTGGCTCAGCGTTGGGGCGGCAAACTGAG 666
QY 652 GTCATGTATACACAGCTTCAAGGCGCTCGAGAGATTGCTGGAGTCTAGTGCAGGTCTGTGAC 711
Db 667 GTCATGTATACACAGCTTCAAGGCACTGGCGCACTCGTGGACTACGTGAGGTCTGTGAC 726
QY 712 TCGGACACAAGGTGGACCCCATGTCACCTGCTGGAGCTGTCGGGTACTTGGACGAGGAC 771
Db 727 TCAGACACAAGACTAGACCCCATGTCACCTGCTGGAGCTTGTGCGAGTGTGGATGAAGAC 786
QY 772 CCCCCTAGAGGCTGTTGGTGGGAGCGTGGGATCTTTAAACCTCTGACACTCTCGGTC 831
Db 787 CCCCCTAGAGGCTGTTGGAGGGGATGTGAGGATCTTTAAACCTCTGACACTCTCGGTC 846
QY 832 AGCTTCTTAAGCAGCTCGATACCTAGCTTGGTGGTGGCTTCAATGTGGAGCGGCTTGTGAG 891
Db 847 AGCTTCTTGAAGCAGCTTTCGATACCTGGGTAGGCTTCAATGTGGAGCGGCTTGTGAG 906
QY 892 TACTTCCACTGTGTATCTTCGATCAGCGGTTCTCTAGGT 930
Db 907 TACTTCCACTGTGTCTCTGATCAGTGTCTCTGGGT 945

RESULT 7
US-10-363-345A-13975
; Sequence 13975, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 13975
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-Island No: 13975
US-10-363-345A-13975

Query Match 39.4%; Score 422; DB 18; Length 662;
Best Local Similarity 77.3%; Pred. No. 2.3e-95;
Matches 512; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

QY 34 GCAGCCCGCGCTGCTCGGCTTCCCGGAGGGTGTGACCATCGCTTCCGCTTGGCTC 93
Db 1 GTAGTTTGTCTGTGTTTCCGTTTGTTCGAGGGGTGTTGATATCCTTTTCTTTT 60
QY 94 ATCTTGGGCTCATGACCTGGGCTTACGCCCGGGGTGCGGCTCGGCTCGGATCGTAC 153
Db 61 ATTTGGGTTTATGATTGGGTTTACGTCGCGGGTGTGTTGGTTTTCGATCGTTAC 120
QY 154 GCGCTCTCGGCTTCCGCTTACGGGCTTCTTCCCTTTCAGCGCACTCTGTGCGCAGAGC 213
```


Qy	95	TCCTGGGCTCATGACCTGGGCTTAGCGCGCGGGGTGCGCTTGGCTTCGATCGTAGC	154
Db	601	TCCTAAACCTCATAAACCTAAACCTTAGCGCGCGGAAATACCGTAAACCTCCGATCGTAGC	542
Qy	155	GCCTCTCGGCTTGGGCTCTACGGGGCTTCTTTTCAGCGCACCTGTGTGGCGCAGAGCC	214
Db	541	ACCTCTAACCTTCGACCTCTACGAAACCTTCTCTTTCAACGCACTTAATACGAAAAACC	482
Qy	215	TCCTGGCTACCTGAGACACCGCGGGGTGGCGCGCGCGCGGGCGGCTTGGATGCGAG	274
Db	481	TCTTCGGCTACTTAAACACCGACGAAATAACGACGACGCGGAAACCGCTAATACAA	422
Qy	275	CCACCGCGCAGTGTGGGCTGACCATCTCCGCTTACGAGGAGACCCCGGTACTGCG	334
Db	421	CCACCGCGCAATATAACGCTTAAACCTCTCCGCTTACCAAAAAAACCCCGGTACTTAC	362
Qy	335	GCCAGTGCCTGGCGTCGCGCGCGCCCTGTCTACCCGCGCGCGGCTGCGGCTCCTCA	394
Db	361	GCCATATCTTAAGTTCGCGCGCCCTTCTATATCCGCGCGCGGCTACGCGTCCCTCA	302
Qy	395	TGCTGGTGGATGGCAACCGCGCCGAGGACCTCTCATCGTTCGACATGTTCCGCGAGGCT	454
Db	301	TAATAATAATAACCAACCGCGCGGAAACCTTCTACATAATCGACATATTCGCGGAACT	242
Qy	455	TCGCTGACGAGGACCCCGCCACGCTACGTGTGGGACGGCAATACACACGCTTGGGAAC	514
Db	241	TCGCTTAACGAAAAACCCGCGCAGTACGTATATAAACGACAACTACCAACCAACCTTAA	182
Qy	515	CCGCGCGCGGGCGGCTGGCGCGGAGCTATCGGGAGGTGAGCGGAGGATCCTG	574
Db	181	CCGCGACGACGACGCGATAAACCGCGGAAACCTATCGAAAAATAAAAAACGAAAAATCCT	122
Qy	575	GCGCGCTGGCAGTGAGCGGCTGTGAGGACTCGCGAGGTGCGTGTGCGTGGCGCAGCGCT	634
Db	121	AACGACTTAACATAAAAAAGCTAATAAAAACTCGCAATACGTATACGTAAACGCAACGCT	62
Qy	635	GGGCGCGCAACGCGAGGTCAATGATACAGGCTTCAAGGCGCTCGGAGATTCGGTGGACT	694
Db	61	AAAAACGACAAACGGGAAATCATATACAAACCTTCAAAACGCTCGAAAAATTCGATAAACT	2
Qy	695	A 695	
Db	1	A 1	

RESULT 10

US-10-363-345A-13974

; Sequence 13974, Application US/10363345A

; Publication No. US20040234960A1

; GENERAL INFORMATION:

; APPLICANT: Alexander Olek

; APPLICANT: Christian Piepenbrock

; APPLICANT: Kurt Berlin

; TITLE OF INVENTION: Method for determining the degree of methylation of defined

; FILE OF INVENTION: cytosines in genomic DNA in the sequence context of 5'-CpG-3'

; FILE REFERENCE: E01/1227

; CURRENT APPLICATION NUMBER: US/10/363,345A

; CURRENT FILING DATE: 2003-03-03

; NUMBER OF SEQ ID NOS: 40712

; SEQ ID NO 13974

; LENGTH: 662

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; OTHER INFORMATION: CpG-island No: 13974

US-10-363-345A-13974

Query Match	36.8%	Score 391.8;	DB 18;	Length 662;
Best Local Similarity	74.7%	Pred. No. 2.4e-88;		
Matches 494;	Conservative	0;	Mismatches 167;	Indels 0; Gaps 0;

Qy	35	CAGCCCGCGCTGTCTCGGCGCTGGCGGAGGGTGCTGACCATCGCCTTGGCCTGCTCA	94
----	----	---	----

Db	2	CAACCTACCGCTACTCCGACCTAAACCCGAAAAATACTAAACCATCGCCTTCGCGCCCTACTCA	61
Qy	95	TCCTGGGGCTCATGACCTGGCGCTACGCGCGCGGGTGCCTGCGCTCCGATCGCTAGC	154
Db	62	TCCTAAACCTCATAACTTAACCTACGCGCGCGGAATACCGCTAACCTCCGATCGTAGC	121
Qy	155	GCCTCTGGCTTCGCGCTTACGCGGCTTCCTTTTCAGCGCACTGTGTGGCGCAGAGCC	214
Db	122	ACCTCTAAACCTTCGACCTTACGAAACCTTCTCTTCAACGCACTTAATACGCAAAAC	181
Qy	215	TCCTCGCGTACTCGAGACCGCGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	274
Db	182	TCCTTCGCGTACTTAAACACCGACGAATAACGACGACGACGCGAATAACCGCTAAATACAA	241
Qy	275	CCACGCGCGCAGTGTGCGCTGACCATCTCCGCTTACACGAGGAGCACCCGCGTACCTGC	334
Db	242	CCACGCGCGCAATATAAGCTAACCATCTCGCTTACCAAAAAACCCCGCTACCTAC	301
Qy	335	GCCAGTGCCTGGCTGCGCGCGCGCGCTGTGTATCCCGCGCGCGCGCTGCGCTCTCA	394
Db	302	GCCAACTACTTAACGCTCCGCGCGCGCTTACTATACCCGCGCGCGGACTACGCGCTCTCA	361
Qy	395	TGCTGTGTGATGGCAACCGCGCCGAGGACCTCTACATGTCGACATGTTTCGCGAGGCTC	454
Db	362	TAATAATAATAAACCAACCGCGCGCGGAACCTTCACTAATCGACATATTCGCGAAATCT	421
Qy	455	TCGCTGACGAGGACCCCGCCACGTACGTGTGGGCGGCAACTACCACGAGCCCTGGGAAC	514
Db	422	TGCTAACGAAAAACCCCGCCAGCTACGTATAAAACGACCAACTACCACCAACCTAAAAAC	481
Qy	515	CCGCGCGCGCGCGCGCGTGGCGCGCGAGCTATCGGAGGTTGAGGCGGAGGATCCTTG	574
Db	482	CCGCGACGACGAACGCGATAAACGCGCGAAACCTATCGAAAAATAAAAAACGAAAAATCCTA	541
Qy	575	GCGCGCTGGCAGTGGAGCGCTGGTGTAGGACTCGCAGGTGCGTGTGCGTGGCGCAGCGCT	634
Db	542	AACGACTAACATAAAAAAGCTATATAAACTCGCAATATCGTATACGTAAACGCAACGCT	601
Qy	635	GGGCGCGCAACGCGAGGTCTATGACACAGCTTCAAGCGCTTCGGAGATTCGGTGGACT	694
Db	602	AAAAACGACAAACGCGAAATCATATACACAACTTCAAAAACGCTCGAAAAATTCGATAAACT	661
Qy	695	A 695	
Db	662	A 662	
RESULT 11			
US-09-918-995-27210			
; Sequence 27210, Application US/09918995			
; Publication No. US20030073623A1			
; GENERAL INFORMATION:			
; APPLICANT: Hyseq, Inc.			
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED			
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES			
; FILE REFERENCE: 20411-756			
; CURRENT APPLICATION NUMBER: US/09/918,995			
; CURRENT FILING DATE: 2001-07-30			
; PRIOR APPLICATION NUMBER: US/09/235,076			
; PRIOR FILING DATE: 1999-01-20			
; NUMBER OF SEQ ID NOS: 38054			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 27210			
; LENGTH: 490			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			
; NAME/KEY: misc_feature			
; LOCATION: (1)...(490)			
; OTHER INFORMATION: n = A,T,C or G			
US-09-918-995-27210			

```

RESULT 11
US-09-918-995-27210
; Sequence 27210, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hvsseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27210
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-27210

```

```
Query Match      31.6%; Score 338.6; DB 10; Length 490;
Best Local Similarity 91.5%; Pred. No. 1.3e-74;
Matches 353; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 544 GCCTATCGGAGGTGGAGCGGAGGATCCTGGGCGGCTGGCAGTGGAGCGCTGGTGGAGG 603
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 GCAGGTAGGAAANNNNNAATTGAGACTTGAGGCACNGTCGGGGAANTCGTGGTGGAG 61

QY 604 ACTCGAGGTGCTGTGCTGGCGGAGCGCTGGGCGGCAAGCGCGAGGTCTATGTACACA 663
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 ACTCGAGGTGCTGTGCTGGCGGAGCGCTGGGCGGCAAGCGCGAGTGTATGTACACA 121

QY 664 GCCTTCAGGCGCTCGGAGATTGCGTGGACTAGCTGCAAGTGTGACTCGGACACAAGG 723
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 GCCTTCAGGCGCTCGGAGATTGCGTGGACTAGCTGCAAGTGTGACTCGGACACAAGG 181

QY 724 TTGGACCCCATGCTGCTGGAGCTCGTGGGCTACTGGACGAGAACCCCGGTAGGG 783
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 TTGGACCCCATGCTGCTGGAGCTCGTGGGCTACTGGACGAGAACCCCGGTAGGG 241

QY 784 GCTGTGTGGGGAAGCTGGGATCTTAAACCTCTGGACTCTCTGGGTGAGTTCCTAAGC 843
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 GCTGTGTGGGGAAGCTGGGATCTTAAACCTCTGGACTCTCTGGGTGAGTTCCTAAGC 301

QY 844 AGCCTGCATACTGGGTGACCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCACTGT 903
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 AGCCTGCATACTGGGTGACCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCACTGT 961

QY 904 GTATCTGTCATCAGCGGTCTCTAGG 929
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 GTATCTGTCATCAGCGGTCTCTAGG 387
```

RESULT 12

```
US-10-295-027-369
; Sequence 369, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
```

```
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 369
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-369
```

```
Query Match      25.5%; Score 272.8; DB 17; Length 1662;
Best Local Similarity 61.0%; Pred. No. 2.6e-58;
Matches 533; Conservative 0; Mismatches 302; Indels 39; Gaps 4;

QY 61 CGGAGGTGCTGACCATGCGCTTCGCCCTGCTATCTCTGGCCTCATGACCTGGCCCTAC 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28 CGTGTGTGGGCACACGCCCTGTTTGGCCCTGGCAGTGTGGGTGGCCTCTGGCAGCCTAT 87

QY 121 GCCGCCGGGGTGGCGCTGGCCTCCGATCGCTACGGCCTCTCTGGCCTTCGGGCTCTACGGG 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 88 GTGAGGGGCTACGAGTTTATCCACACGAAAGACACTACCTGTCTCTGGCCTGTACGGC 147

QY 181 GCCTTCTCTTTCAGCGCACCTGGTGGCGCAGAGCCTCTTTCGCTACCTGGAGCACCGGCG 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 148 GCCATCTCTGGGCTGACCTGCTCATTTACAGAGCCTTTTGGCCTTCTTGGAGCACCGGGCG 207

QY 241 GTGCGGGGGGGCGCGCGGGGCGCTGGATGACGACACCGCGC---GCAGTGTGGCGCTG 297
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 208 ATGCGAGCTGCGCGCCAGGCCCTGAAGCTGCCCTCCCGCGCGGGGCTCGGTGGCACTG 267

QY 298 ACCATCTCCGCTACACGAGGAGCCCGCGTACCTGGCGCACAGTGGCTGGCGCTCGCGCCGC 357
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 268 TGCATTTGGCGATACACGAGGAGCCCTGACTACTTTGGGCAAGTGGCTTCGCTTCGGCCAG 327

QY 358 GCCCTGCTGTACCCGCGCGCGGCTGCGCTCTCTCATGTGTGGATGGCAACCGCGCC 417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 328 CGCATCTCTTCCCTGAC-----CTCAAGTGTGTCTATGGTGGATGGCAACCGCCAG 381

QY 418 GAGGACCTCTACATGTTGCGATGTTTCGGCGAGTCTTCGCTGACGAGGACCCCGCCAG 477
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 382 GAGGACGCTACATGCTGGACATCTTCCAGAGTGTGGCGGCAACGAGCACCGCGCGC 441

QY 478 TACGTGTGGAGCGCAACTACACACGCTCTGGGAAACCCCGCGCGGGCGCGCTGGGC 537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 TTCTTTGTGGCGCAGCACTTCCA-----TGAGGACAGGCGAGGTTGAG 486

QY 538 GCCGAGCCTATCGGAGGTGGAGCGGAGATCCTGGCGGCTGGCAGTGGAGCGCGCTG 597
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 487 ACGGAGCCAGCCTGCGAGGAGGCGATGGACCGT-----GTGCGGATGTG 531

QY 598 GTGAGGACTCGCAGTGGTGTGGCGCAGCGCTGGGCGCGCAAGCGCGAGGTCTATG 657
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 532 GTGCGGGCCAGCACCTTTCTGTGATCATGCAAGATGGGAGGCAAGCGCGAGGTCTATG 591

QY 658 TACACAGCCTTCAAGCGCTCGGAGATTTCGTGTGACTACGTGACAGTCTGTGACTCGGAC 717
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 592 TACACGCGCTTCAAGGCGCTCGGCGATTCGGTGTGACTACATCCAGGTTGGGACTCTGAC 651

QY 718 ACAAGTTTGGACCCCATGGCACTGTCTGAGCTGTGCGGGTACTGGAAGAGGACCCCGCG 777
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 652 ACTGTGTGGATCCAGCCTGCACCATCGAGATGCTTCGAGTCTCTGGAGGAGATCCCCAA 711

QY 778 GTAGGGGCTGTTGGTGGGAGCGTGGGATCCTTAAACCTCTGGACTCTCTGGGTGAGCTTC 837
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 712 GTAGGGGAGTTCGGGGGAGATGTCCAGATCCTCAACAAGTACGACTCATGATTTCTTC 771

QY 838 CTAAGCAGCCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTTC 897
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 772 CTGAGCAGCGTGGGTACTGTGATGCGCTTCAACGTGGAGCGGGCTGCGCACTCTACTTT 831

QY 898 CACTGTGTATCTCTGCATCAGCGGTTCTTAGGTA 931
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 832 GGCTGTGTGAGTGTATTAGTGGGCCCTTGGCA 865
```

```
RESULT 13
US-10-188-832-136
; Sequence 136, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Azi, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 136
; TYPE: DNA
; LENGTH: 1662
; ORGANISM: Homo sapiens
US-10-188-832-136

Query Match      25.5%; Score 272.8; DB 17; Length 1662;
Best Local Similarity 61.0%; Pred. No. 2.6e-58; Indels 39; Gaps 4;
Matches 533; Conservative 0; Mismatches 302;

QY 61 CGAGGGGTGCTGACCATCGCCCTTCGCGCTGCTCATCTCGGGCTCATGACCTGGGCGCTAC 120
DB 28 CGTGTGGTGGCCACGAGCTGTTGCCCTGGCAGTGTGGTGGCATCTCGGCGCTAT 87
QY 121 GCGCGGGGTGCGGTGGCTCCGATCGCTAGGCGCTCTGGCGCTTGGCGCTCTACGGG 180
DB 88 GTGAGGGGTACAGTTTATCCACACGGAAGACATACCTGTCTCTCGGCGCTGTACGGC 147
QY 181 GCCTTCTTTACGCGCACTGTGGCGAGAGCTCTTCGCGTACTCTGAGCAGCAGCGCGG 240
DB 148 GCATCTCTGGCGCTGACCTGCTCATTTACAGAGCTTTTGGCTTCTTGAGCAGCAGCGCGC 207
QY 241 GTGGCGGGCGCGCGGGCGCGTGGATGAGCAGCAGCGCGC---GCAAGTGTGGCGGTG 297
DB 208 ATGCGAGCTGCGCGCGCGCGCTGAAGCTGCCCTCCCGCGCGGGGCTCGGTGGCAGTG 267
QY 298 ACCATCTCGCGCTACAGAGAGACCCCGGTACTGTGCGCAAGTGTGCGCTGCGCGCGCG 357
DB 268 TGCAATGCGCGTACAGAGAGACCTGACTACTTGGCAAGTGTGCGCTGCGCTCGCGCGCAG 327
QY 358 GCGCTGCTGTACCGCGCGCGCGGTGCGCGTCTCATGCTGCTGATGGTGGTGGTGGTGGT 417
DB 328 CGCATCTCTCTCCCTGAC-----CTCAAGGTGGTCTGATGGTGGTGGTGGTGGTGGT 381
QY 418 GAGGACCTCTACATGCTGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCAGG 477
DB 382 GAGGACGCTACATGCTGACATCTTCACAGAGTGTGCTGGCGCGGACCGAGCAGCGCGC 441
QY 478 TACGTGTGGGACGGCAACTACACAGCGCTTGGGAACCGCGCGCGCGCGCGGTGGGCG 537
DB 442 TTTCTTGTGTGCGCGAGCAACTTCCA-----TGAGGAGGCGAGGCTGAG 486
QY 538 GCGGAGCGCTATCGGAGGTGAGGCGGAGGATCTTGGCGCGCTGGCAGTGGAGCGGTG 597
DB 487 ACGGAGCGCGAGCTTCAGAGGGGATGAGCCGT-----GTGCGGATGTG 531

RESULT 14
US-10-309-560-9
; Sequence 9, Application US/10309560
; Publication No. US20030235893A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: KUMARI, KSHAMA
; TITLE OF INVENTION: HYALURONAN SYNTHASES AND METHODS OF MAKING AND USING SAME
; FILE REFERENCE: 35541.082
; CURRENT APPLICATION NUMBER: US/10/309,560
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 60/336,105
; PRIOR FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Xenopus laevis
US-10-309-560-9

Query Match      23.2%; Score 248.4; DB 17; Length 1767;
Best Local Similarity 56.3%; Pred. No. 3e-52;
Matches 490; Conservative 0; Mismatches 371; Indels 9; Gaps 1;

QY 62 GGAGGGTCTGACCATCGCCCTTCGCGCTGCTCATCTCGGGCTCATGACCTGGGCTAGC 121
DB 80 GGAGGATTAATTATTTATTTCTTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 139
QY 122 CGCGCGGGTGGCGGTGGCGCTCCGATCGCTACGGCTCTCGCGCTCTCGCGCTCTACGGGG 181
DB 140 TGGCAGAGTTCAGGTCTCTCAACATGAAGCAATTTCTCTCTCTCTCTCTCTCTCTCT 199
QY 182 CTTCTCTTTACGGCAGCCTGTGGCGAGAGCTCTTCGCGTACTCGAGCAGCGCGGG 241
DB 200 TTGCAATCTTTCTCACCTGATGACAGAGCTCTTTTGGCTTCTGAGATACGAGG 259
QY 242 TGGCGCGCGCGCGGGGGCGCTGGATGAGCAGCCACCGCGAGTGTGGCGCTGACCA 301
DB 260 TAAATAAGAGTGA-----GCTTCTTGGAGCTTTTAAAGAGCAGTGGCTCTGACCA 310
QY 302 TCTCGCGCTACAGGAGGACCCCGGTACTCGCGCAGTGGCGCTGGCGCTCCCGCGCGCC 361
DB 311 TTGCTGGGTATCAGGAGAACCTCGAGTACTGATAAAGTCTTGGTGAATCTCTCAAGTATG 370
QY 362 TGCTGTATCCCGCGCGCGGTGCGCGTCTCTCATGTGGTGGTGGATGGCAACCGCGCGGAGG 421
DB 371 TGAATAACCCCAAGATAAACTCAAGATCAATTTTGGTCAATCGATGGGAAACACAGAGGATG 430
```

```
Qy 422 ACCTTACATGCTGACATGTTCCCGAGGTTCTTCGCTGACGAGACCCCGCCAGGTACG 481
Db 431 ATGCTACATGATGAGATGTTTCAAGACGTTTCCACGGTGAAGATGTAGGCACCTACG 490
Qy 482 TGTGGAGCGGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGCGGCTGGGCGCGG 541
Db 491 TATGGAAGGGAATTTACCACTGTTAAAGCCTGAGGAGACCAATTAAGGGATCTGTGC 550
Qy 542 GAGCTATCGGAGAGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGCGGCTGGTGA 601
Db 551 CTGAGGTTTCTAAGCCCTTGAATGAAGATGAAGGTATCAATATGTTGGAAGAACTTTGTA 610
Qy 602 GGACTCGAGGTGCTGTGCGTGGCGGACGCTGGGGCGGCAAGCGCGAGGTCTATGTACA 661
Db 611 GAAACAAGAGATGTTGTGTGTCATCATGCAACAGTGGGGCGGAAAAAGAGAGGTCTATGTACA 670
Qy 662 CAGCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAGTCTGTGACTCGGACACAA 721
Db 671 CAGCAATTCAGGCCATTTGGGACTTCTGTGGACTATGTACAGGTCTGTGACTCGGACACA 730
Qy 722 GGTGGACCCCATGSCATGCTGGAGCTGCTGGCGGTACTGGAAGAGACCCCGGGTAG 781
Db 731 AACTGGATGAATGTGCAACAGTGGAAATGGTGAAGGTTCTGGAATCCAAATGACATGTACG 790
Qy 782 GGGCTGTTGGTGGGAGCTGGGATCCTTAACCTCTGGACTCTGCTGGGTCAAGTTCCTTAA 841
Db 791 GCGCAGTGGGAGGAGCGTTGCGCATTTCTGAACCTTATGATTCCTTCAATTAGTTTCAATGA 850
Qy 842 GCAGCTCGGATACCTGGGTAGCGTTCAATGTGGAGCGGCTTGTGAGAGTACTTCCACT 901
Db 851 GCAGCTCGGTACTGGATGGGTTTAACTGGAGAGGCGCTGCGAGGCTTCTTACTTCGACT 910
Qy 902 GTGATCTGTCATCAGCGGTTCTCTAGGTA 931
Db 911 CGGTGCTCTGTAAAGTGGACCTCTGGGA 940
```

RESULT 15

```
US-09-902-939-3
; Sequence 3, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHaza
; APPLICANT: Weilliam Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/OH020-US0
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: mouse
US-09-902-939-3
```

Query Match 22.1%; Score 236.2; DB 10; Length 1665;

Best Local Similarity 58.4%; Pred. No. 3.3e-49;

Matches 512; Conservative 0; Mismatches 323; Indels 42; Gaps 4;

```
Qy 61 CGGAGGGTGTGACCATCGCTTGGCCCTGCTCATCTGGGCGCTCATGACCTGGGCTTAC 120
Db 28 CGTGTGGTGGCACCAGTCTGCTTTGGCCCTGGTAGTGTGGGAGGCATCTCGCGGCTTAT 87
Qy 121 GCCCGCGGGGTGCGCGCTCGGCCTCGCATCGCTACGGGCTTCTGGGCTTCTAGCGG 180
Db 88 GTGACAGGCTACCACTTTATCCACACAGAAAGCACTACCTGTCTTTGGCTCTACGGT 147
Qy 181 GCCTTCTCTTCAGCGCACCTGGTGGCGGAGAGCCTTCTTGGGTACTCTGGAGCAACGGCGG 240
Db 148 GCCATCTGGGTCTACATCTGCTCATCCAGAGCCTGTTTGGCTTCTCTGGAGCACCGTCA 207
```

```
Qy 241 GTGGCGGCGGCGCGCG-----GGGGCGGCTGATGCAGCCACCGCGCGCAGTGTGGCG 294
Db 208 ATGGCGAGGCGAGGGCGGCCCTCAAGCTGCACCTGCTCCCAGAGGTGCGTTCAGTGGCA 267
Qy 295 CTGACCATCTCCGCTTACCAGGAGGACCCCGCTACTCTGCGCCAGTGCCTCGCTCGGCC 354
Db 268 CTCTGCATTGCTGCTTACCAAGAGGACCCCGAATACCTGCGCAAGTGCCTTCGCTCAGCT 327
Qy 355 CGCCCTCTGCTGTATCCCGCGCGCGGCTGGCGCTCTCATGTTGGTGGATGGCAACGCC 414
Db 328 CAGCGCATTTGCTTTTCCAAAC-----CTCAAGGTGGTTCATGTTAGTGGATGGCAATGCG 381
Qy 415 GCCGAGGACCTCTACATGTTGACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCC 474
Db 382 CAGGAAGATACCTACATGTTGGACATCTTTCATGAGGTGCTGGTGGCACTGAGCAAGCT 441
Qy 475 ACGTACGTGTGGGACGGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGCGCTG 534
Db 442 GGCTTCTTTGTTGTTGGCTGAGCAATTTCCA-----TGAGGCGGGTGAAAGGA 486
Qy 535 GGCCCGGAGGCTATCGGAGGTGGAGCGGAGGATCTCTGGGCGGCTGGCAGTGGAGCGG 594
Db 487 GAGACAGAGGCGCACCTGCGAGGAAGGCAATGAGCGT-----GTGCGAGCT 531
Qy 595 CTGTTGAGGACTCGCAGTGTGCTGCTGCGCAGCGCTGGGCGCGGCAAGCGGAGGTC 654
Db 532 GTGTTGTTGGGCGACACCTTCTCATGCAATATGAGAGTGGGGGGGCAAGCGTGAAGTC 591
Qy 655 ATGTACACAGCCTTCAAGGGCGCTCGGAGATTCGGTGGACTACGTCGAGGTCTGTGACTCG 714
Db 592 ATGTACACTGCTTCAAGGCCCTTGGCAACTCTAGTGGACTACATCCAGGTGTGTGACTCT 651
Qy 715 GACAAGAGTTGGACCCCATGGGCACTGCTGGAGCTCGTGGGGTACTGGAGGAGACCCC 774
Db 652 GACACTGTGCTGGACCCAGCCTGCACCATTTGAGATGCTTCGAGTCTTGGAGAAAGATCCC 711
Qy 775 CGGTTAGGGGCTGTTGGTGGGAGCTGGGATCCTTAACCTCTGGACTCTGGGTGAGC 834
Db 712 CAAGTAGAGGTGTTGGAGGAGATGTCAAATCCTCAAGATATGATTCATGATCTCC 771
Qy 835 TTCTTAAGCAGCCTTCGATATCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTTCAGAGTAC 894
Db 772 TTCTCTGAGCAGTGTGAGGTACTGGATGGCTTTCAACGCTGGAGCGGCGCTGCCAGTCTTAC 831
Qy 895 TTCCACTGTGTATCTTCATCGATCAGCGGTTCTCTAGGTA 931
Db 832 TTTGGCTGTGTGCAATGTATTTAGTGGGCTTTTGGGCA 868
```

Search completed: March 13, 2005, 17:09:55

Job time : 758.899 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:01:16 ; Search time 210.96 Seconds
(without alignments)
8307.054 Million cell updates/sec

Title: US-10-672-399-3
Perfect score: 1071
Sequence: 1 atgagacagcaggacgccc.....ccggaccatcggtgatga 1071

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	927.4	86.6	2116	3	US-08-635-552A-1 Sequence 1, Appli
2	924.2	86.3	2087	4	US-09-949-016-5659 Sequence 5659, Ap
3	915	85.4	2117	3	US-09-155-768-1 Sequence 1, Appli
4	697.2	65.1	14862	4	US-09-949-016-17401 Sequence 17401, A
5	396.4	37.0	601	4	US-09-949-016-201775 Sequence 201775,
6	272.8	25.5	1662	4	US-09-949-016-3728 Sequence 3728, Ap
7	236.2	22.1	1665	4	US-08-812-008-31 Sequence 31, Appli
8	230.8	21.5	601	4	US-09-949-016-201780 Sequence 201780,
9	176.2	16.5	2947	4	US-08-675-499A-1 Sequence 1, Appli
10	176.2	16.5	2947	4	US-08-812-008-1 Sequence 1, Appli
11	167.6	15.6	9871	4	US-09-949-016-15470 Sequence 15470, A
12	148.8	13.9	3003	4	US-09-949-016-737 Sequence 737, App
13	137.4	12.8	601	4	US-09-949-016-132199 Sequence 132199,
14	82.8	7.7	6645	4	US-09-902-540-8819 Sequence 8819, Ap
15	82.8	7.7	10210	4	US-09-902-540-938 Sequence 938, App
16	69.2	6.5	1320	4	US-07-724-797-85 Sequence 85, Appl
17	66.6	6.2	3957	4	US-10-237-551-193 Sequence 193, App
18	66.6	6.2	154746	4	US-09-827-688-8 Sequence 8, Appli
19	66.6	6.2	154746	4	US-09-827-688-8 Sequence 8, Appli
20	66.2	6.2	1614	4	US-09-616-289-45 Sequence 45, Appl
21	66.2	6.2	12425	4	US-09-616-289-50 Sequence 50, Appl
22	62.2	5.8	32176	4	US-09-949-016-12479 Sequence 12479, A
23	62	5.8	1292	3	US-08-483-533-37 Sequence 37, Appl
24	62	5.8	1292	3	US-09-283-471A-37 Sequence 37, Appl
25	60.6	5.7	4257	2	US-08-690-473-1 Sequence 1, Appli
26	60.6	5.7	4257	3	US-09-259-821A-1 Sequence 1, Appli
27	60.6	5.7	4257	3	US-08-843-659-1 Sequence 1, Appli

28	60.6	5.7	4257	4	US-09-825-288A-1 Sequence 1, Appli
29	60.6	5.7	12001	1	US-08-458-568A-11 Sequence 11, Appl
30	60.2	5.6	2900	1	US-08-034-650-9 Sequence 9, Appli
31	60.2	5.6	2900	1	US-08-449-015-9 Sequence 9, Appli
32	59.6	5.6	1647	4	US-09-902-540-4211 Sequence 4211, Ap
33	59.6	5.6	21330	4	US-09-902-540-1209 Sequence 1209, Ap
34	59.4	5.5	1926	3	US-09-249-585A-4 Sequence 4, Appli
35	59.4	5.5	1931	2	US-09-130-114-2 Sequence 2, Appli
36	59.2	5.5	441529	3	US-09-103-840A-1 Sequence 1, Appli
37	59.2	5.5	1051	2	US-08-865-273-1 Sequence 1, Appli
38	59.2	5.5	1051	3	US-09-385-174-1 Sequence 1, Appli
39	59.2	5.5	1182	4	US-09-673-198-31 Sequence 31, Appl
40	59.2	5.5	1361	4	US-09-614-912-37 Sequence 37, Appl
41	59	5.5	4403765	3	US-09-103-840A-2 Sequence 2, Appli
42	58.4	5.5	5808	4	US-09-902-540-794 Sequence 794, App
43	58.2	5.4	967	4	US-09-902-540-7886 Sequence 7886, Ap
44	57.8	5.4	8604	4	US-09-902-540-5331 Sequence 5331, Ap
45	57.8	5.4	33399	4	US-09-902-540-1260 Sequence 1260, Ap

ALIGNMENTS

RESULT 1
US-08-635-552A-1
; Sequence 1, Application US/08635552A
; Patent No. 6423514
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; TITLE OF INVENTION: Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1769
US-08-635-552A-1

Query Match 86.6%; Score 927.4; DB 3; Length 2116;
Best Local Similarity 99.9%; Pred. No. 3.2e-175;
Matches 928; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAGACAGCAGGACGCCCAAGCCCACTCTCTGACGCCCGCTGCTCGGCTGCGC 60
|||||

```
Db 36 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGACGCGCGCGCTGCTCGGCGCTGGCC 95
Qy 61 CGGAGGGTGTGACCATGCGCTTGGCCCTGTCTATCTCTGAGGCTCATGACCTTGGGCTTAC 120
Db 96 CGGAGGGTGTGACCATGCGCTTGGCCCTGTCTATCTCTGAGGCTCATGACCTTGGGCTTAC 155
Qy 121 GCGCGCGGGGTGCGGCTGCGCTCGCATCGCTACGCGGCTCTGGGCTTCTGAGGCTTACGGG 180
Db 156 GCGCGCGGGGTGCGGCTGCGCTCGCATCGCTACGCGGCTCTGGGCTTCTGAGGCTTACGGG 215
Qy 181 GCCTTCTCTTTCAGCGCACCTGTGTGCGCAGAGCTCTTTCGGGTACTCTGGAGCACCGGCGG 240
Db 216 GCCTTCTCTTTCAGCGCACCTGTGTGCGCAGAGCTCTTTCGGGTACTCTGGAGCACCGGCGG 275
Qy 241 GTGGCGGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
Db 276 GTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 335
Qy 301 ATCTCCGCTTACCAAGGAGGACCGCGCTGTGATGAGCCACCGCGCGAGTGTGGCGCTTAC 360
Db 336 ATCTCCGCTTACCAAGGAGGACCGCGCTGTGATGAGCCACCGCGCGAGTGTGGCGCTTAC 395
Qy 361 CTGCTGTATACCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
Db 396 CTGCTGTATACCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 455
Qy 421 GACCTCTACATGTGTGAGACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCGCACGTAC 480
Db 456 GACCTCTACATGTGTGAGACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCGCACGTAC 515
Qy 481 GTGTGGGACCGGCAACTACCAACCGCTGGGAAACCGCGCGGCGGCGGCGGCGGCGGCGGCGG 540
Db 516 GTGTGGGACCGGCAACTACCAACCGCTGGGAAACCGCGCGGCGGCGGCGGCGGCGGCGGCGG 575
Qy 541 GTGTGGGACCGGCAACTACCAACCGCTGGGAAACCGCGCGGCGGCGGCGGCGGCGGCGGCGG 600
Db 576 GTGTGGGACCGGCAACTACCAACCGCTGGGAAACCGCGCGGCGGCGGCGGCGGCGGCGGCGG 635
Qy 601 AGGACTCGCAGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 660
Db 636 AGGACTCGCAGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 695
Qy 661 ACAGCTTCAAGGCGCTTCGAGATTCGGTGGATTCGGTGGATTCGGTGGATTCGGTGGATTCGG 720
Db 696 ACAGCTTCAAGGCGCTTCGAGATTCGGTGGATTCGGTGGATTCGGTGGATTCGGTGGATTCGG 755
Qy 721 AGGTTGAGACCCCATGCGACCTGCTGGAGCTGCTGGGCTGCTGGGCTGCTGGGCTGCTGGGCT 780
Db 756 AGGTTGAGACCCCATGCGACCTGCTGGAGCTGCTGGGCTGCTGGGCTGCTGGGCTGCTGGGCT 815
Qy 781 GGGGCTGTGTGGGAGCTGCGGATCTTAACCTCTGGAATCTCTGGGCTGAGCTTCTCTTA 840
Db 816 GGGGCTGTGTGGGAGCTGCGGATCTTAACCTCTGGAATCTCTGGGCTGAGCTTCTCTTA 875
Qy 841 AGCAGCTCCGATCTGAGTACCTTCAATCTGGAGCGGCGGCTTCTGAGGCTACTTCCAC 900
Db 876 AGCAGCTCCGATCTGAGTACCTTCAATCTGGAGCGGCGGCTTCTGAGGCTACTTCCAC 935
Qy 901 TGTGTATCTTCATCAGCGGTTCTCTAGG 929
Db 936 TGTGTATCTTCATCAGCGGTTCTCTAGG 964
```

RESULT 2
US-09-949-016-5659
; Sequence 5659, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016

```
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5659  
; LENGTH: 2087  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-5659  
  
Query Match 86.3%; Score 924.2; DB 4; Length 2087;  
Best Local Similarity 99.7%; Pred. No. 1.4e-174;  
Matches 926; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGACGCGCGCGCTGCTCGGCGCTGGCC 60  
Db 36 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGACGCGCTGCTCGGCGCTGGCC 95  
Qy 61 CGGAGGGTGTGACCATGCGCTTGGCCCTGTCTATCTCTGAGGCTCATGACCTTGGGCTTAC 120  
Db 96 CGGAGGGTGTGACCATGCGCTTGGCCCTGTCTATCTCTGAGGCTCATGACCTTGGGCTTAC 155  
Qy 121 GCGCGCGGGGTGCGGCTGCGCTCGCATCGCTACGCGGCTCTGGGCTTCTGAGGCTTACGGG 180  
Db 156 GCGCGCGGGGTGCGGCTGCGCTCGCATCGCTACGCGGCTCTGGGCTTCTGAGGCTTACGGG 215  
Qy 181 GCCTTCTCTTTCAGCGCACCTGTGTGCGCAGAGCTCTTTCGGGTACTCTGGAGCACCGGCGG 240  
Db 216 GCCTTCTCTTTCAGCGCACCTGTGTGCGCAGAGCTCTTTCGGGTACTCTGGAGCACCGGCGG 275  
Qy 241 GTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300  
Db 276 GTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 335  
Qy 301 ATCTCCGCTTACCAAGGAGGACCGCGCTGTGATGAGCCACCGCGCGAGTGTGGCGCTTAC 360  
Db 336 ATCTCCGCTTACCAAGGAGGACCGCGCTGTGATGAGCCACCGCGCGAGTGTGGCGCTTAC 395  
Qy 361 CTGCTGTATACCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420  
Db 396 CTGCTGTATACCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 455  
Qy 421 GACCTCTACATGTGTGAGACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCGCACGTAC 480  
Db 456 GACCTCTACATGTGTGAGACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCGCACGTAC 515  
Qy 481 GTGTGGGACCGGCAACTACCAACCGCTGGGAAACCGCGCGGCGGCGGCGGCGGCGGCGGCGG 540  
Db 516 GTGTGGGACCGGCAACTACCAACCGCTGGGAAACCGCGCGGCGGCGGCGGCGGCGGCGGCGG 575  
Qy 541 GTGTGGGACCGGCAACTACCAACCGCTGGGAAACCGCGCGGCGGCGGCGGCGGCGGCGGCGG 600  
Db 576 GTGTGGGACCGGCAACTACCAACCGCTGGGAAACCGCGCGGCGGCGGCGGCGGCGGCGGCGG 635  
Qy 601 AGGACTCGCAGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 660  
Db 636 AGGACTCGCAGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 695  
Qy 661 ACAGCTTCAAGGCGCTTCGAGATTCGGTGGATTCGGTGGATTCGGTGGATTCGGTGGATTCGG 720  
Db 696 ACAGCTTCAAGGCGCTTCGAGATTCGGTGGATTCGGTGGATTCGGTGGATTCGGTGGATTCGG 755  
Qy 721 AGGTTGAGACCCCATGCGACCTGCTGGAGCTGCTGGGCTGCTGGGCTGCTGGGCTGCTGGGCT 780  
Db 756 AGGTTGAGACCCCATGCGACCTGCTGGAGCTGCTGGGCTGCTGGGCTGCTGGGCTGCTGGGCT 815  
Qy 781 GGGGCTGTGTGGGAGCTGCGGATCTTAACCTCTGGAATCTCTGGGCTGAGCTTCTCTTA 840  
Db 816 GGGGCTGTGTGGGAGCTGCGGATCTTAACCTCTGGAATCTCTGGGCTGAGCTTCTCTTA 840  
Qy 841 AGCAGCTCCGATCTGAGTACCTTCAATCTGGAGCGGCGGCTTCTGAGGCTACTTCCAC 840  
Db 876 AGCAGCTCCGATCTGAGTACCTTCAATCTGGAGCGGCGGCTTCTGAGGCTACTTCCAC 840  
Qy 901 TGTGTATCTTCATCAGCGGTTCTCTAGG 929  
Db 936 TGTGTATCTTCATCAGCGGTTCTCTAGG 964
```


Db 816 GGGGCTGTGGTGGGATGTGGGATCTTAACCTCTTGGACTCTGGACTCTGGGTGACGTTCTCTA 875
Qy 841 AGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGACAGCTACTTCCAC 900
Db 876 AGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGACAGCTACTTCCAC 935
Qy 901 TGTGTATCTGCATCAGCGGTTCTCTAGG 929
Db 936 TGTGTATCTGCATCAGCGGTTCTCTAGG 964

RESULT 3

US-09-155-768-1

; Sequence 1, Application US/09155768A

; Patent No. 6162908

; GENERAL INFORMATION:

; APPLICANT: SEIKAGAKU CORPORATION

; TITLE OF INVENTION: POLYPEPTIDE OF HYALURONAN SYNTHASE OF

; TITLE OF INVENTION: HUMAN-ORIGIN AND DNA ENCODING THE SAME

; FILE REFERENCE: TOYAM37.001APC

; CURRENT APPLICATION NUMBER: US/09/155,768A

; EARLIER FILING DATE: 1998-10-05

; EARLIER APPLICATION NUMBER: JP 8-084326

; EARLIER FILING DATE: 1996-04-05

; EARLIER APPLICATION NUMBER: JP 8-109663

; EARLIER FILING DATE: 1996-04-30

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 1

; LENGTH: 2117

; TYPE: DNA

; ORGANISM: HUMAN

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (149)...(1777)

US-09-155-768-1

Query Match

Best Local Similarity 85.4%; Score 915; DB 3; Length 2117;

Matches 918; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 7 CAGCAGGACGGCCCAAGCCCACTCTCGAGCCCGCGCTGCTCGGGCTGGCCCGGAGG 66
Db 50 CAGCAGGACGGCCCAAGCCCACTCTCGAGCCCGCGCTGCTCGGGCTGGCCCGGAGG 109
Qy 67 GTGCTGACCATCGCTTCGCCCTGCTCATCTGCGGCTCATGACCTGGGCTTACGCCGCC 126
Db 110 GTGCTGACCATCGCTTCGCCCTGCTCATCTGCGGCTCATGACCTGGGCTTACGCCGCC 169
Qy 127 GGGGTGCGGCTGGCTTCGGATCGCTACGCGCTCTCTGGGCTTCTAGGGGCTTC 186
Db 170 GGGGTGCGGCTGGCTTCGGATCGCTACGCGCTCTCTGGGCTTCTAGGGGCTTC 229
Qy 187 CTTTCAGGCGACCTGGTGGCGAGAGCCTCTTGGGCTACTGGAGACCGGGGGTGGG 246
Db 230 CTTTCAGGCGACCTGGTGGCGAGAGCCTCTTGGGCTACTGGAGACCGGGGGTGGG 289
Qy 247 GGGGGGGCGGGGGCGGCTGATCAGCCACCGCGCGAGTGTGGCGCTGACCATCTCC 306
Db 290 GGGGGGGCGGGGGCGGCTGATCAGCCACCGCGCGAGTGTGGCGCTGACCATCTCC 349
Qy 307 GCCTACCAAGGAGACCGCGGCTGCTGCGGCTGCGGCTGCGGCGGCTGCTG 366
Db 350 GCCTACCAAGGAGACCGCGGCTGCTGCGGCTGCGGCTGCGGCGGCTGCTG 409
Qy 367 TACCGCGCGGCGGCTGGGCTCTCTGAGTGTGGATGGCAACCGCGCGGAGACCTC 426
Db 410 TACCGCGCGGCGGCTGGGCTCTCTGAGTGTGGATGGCAACCGCGCGGAGACCTC 469
Qy 427 TACATGTGCGACATGCTTCGGGAGGTCTTCCGTGACGAGGACCGCGGCTTACGTGTGG 486
Db 470 TACATGTGCGACATGTTCCGGGAGGTCTTCCGTGACGAGGACCGCGGCTTACGTGTGG 529

Qy 487 GACGCAATACCACTAGCCCTTGGAAACCCCGCGCGCGCGCTGGGCGCGGAGCC 546
Db 530 GACGCAATACCACTAGCCCTTGGAAACCCCGCGCGCGCGCTGGGCGCGGAGCC 589
Qy 547 TATCGGAGGTGGAGCGGAGATCCTGGGCGGCTGGCAGTGGAGCGCTGTTGAGGACT 606
Db 590 TATCGGAGGTGGAGCGGAGATCCTGGGCGGCTGGCAGTGGAGCGCTGTTGAGGACT 649
Qy 607 GCGAGGTGCGTGTGCGTGGCGCAGCGCTGGGCGGCAAGCGCGAGGTCTATGTACACAGCC 666
Db 650 GCGAGGTGCGTGTGCGTGGCGCAGCGCTGGGCGGCAAGCGCGAGGTCTATGTACACAGCC 709
Qy 667 TTCAAGGCGCTCGGAGATTCGGTGAATCTGTCAGGCTTGTGACTCGGACACAAAGTTG 726
Db 710 TTCAAGGCGCTCGGAGATTCGGTGAATCTGTCAGGCTTGTGACTCGGACACAAAGTTG 769
Qy 727 GACCCCATGCGCACTGCTGGAGCTCGTGGGCTACTGGAGCAGGAGCCCGGGTACGGGCT 786
Db 770 GACCCCATGCGCACTGCTGGAGCTCGTGGGCTACTGGAGCAGGAGCCCGGGTACGGGCT 829
Qy 787 GTTGTGGGAGCGTGGCGATCCTTAACCTCTGGACTCTGGGTGAGTCTTCTAAGCAGC 846
Db 830 GTTGTGGGAGTGGCGATCCTTAACCTCTGGACTCTGGGTGAGTCTTCTAAGCAGC 889
Qy 847 CTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCACTGTGA 906
Db 890 CTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCACTGTGA 949
Qy 907 TCCTGCATCAGCGGTTCTCTAGG 929
Db 950 TCCTGCATCAGCGGTTCTCTAGG 972

RESULT 4

US-09-949-016-17401

; Sequence 17401, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CLO01307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17401

; LENGTH: 14862

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-17401

Query Match

Best Local Similarity 65.1%; Score 697.2; DB 4; Length 14862;

Matches 699; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 7 CAGCAGGACGGCCCAAGCCCACTCTCGAGCCCGCGCTGCTCGGGCTGGCCCGGAGG 66
Db 6073 CAGCAGGACGGCCCAAGCCCACTCTCGAGCCCGCGCTGCTCGGGCTGGCCCGGAGG 6132
Qy 67 GTGCTGACCATCGCTTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCTTACGCCGCC 126
Db 6133 GTGCTGACCATCGCTTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCTTACGCCGCC 6192
Qy 127 GGGGTGCGGCTGGGCTCGGATCGCTAGGCGCTTCTGGGCTTCTAGGGGCTTTC 186
Db 6193 GGGGTGCGGCTGGGCTCGGATCGCTAGGCGCTTCTGGGCTTCTAGGGGCTTTC 6252

```
Qy 187 CTTTCAGCGCACCTGCTGGCGCAGAGCCTCTTCGGGTACCTGGAGCACCGGGGGTGGCG 246
Db 6253 CTTTCAGCGCACCTGCTGGCGCAGAGCCTCTTCGGGTACCTGGAGCACCGGGGGTGGCG 6312
Qy 247 GCGGGCGCGCGGGGCGCTGATGACGACCGCGCAGTGTGGCGCTGACCATCTCC 306
Db 6313 GCGGGCGCGCGGGGCGCTGATGACGACCGCGCAGTGTGGCGCTGACCATCTCC 6372
Qy 307 GCCTACGAGGAGACCCCGCTACTGCGCCAGTGCCTGGCGTCCGCCCGCCCTGCTG 366
Db 6373 GCCTACGAGGAGACCCCGCTACTGCGCCAGTGCCTGGCGTCCGCCCGCCCTGCTG 6432
Qy 367 TACCGCGCGCGCGCTCGGCTCTCATGTGTGGATGGCAACCGCGCGCGAGACCTC 426
Db 6433 TACCGCGCGCGCGCTCGGCTCTCATGTGTGGATGGCAACCGCGCGCGAGACCTC 6492
Qy 427 TACATGTCGACATGTTCCGGGAGGTCTTCGTGACGAGGACCCCGCCACGTACTGTGG 486
Db 6493 TACATGTCGACATGTTCCGGGAGGTCTTCGTGACGAGGACCCCGCCACGTACTGTGG 6552
Qy 487 GACGCCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGGTGGCGCCCGAGCC 546
Db 6553 GACGCCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGGTGGCGCCCGAGCC 6612
Qy 547 TATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGGCGCTGTGAGACT 606
Db 6613 TATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGGCGCTGTGAGACT 6672
Qy 607 CGCAGTGCCTGTGCTGGCGCAGCGCTGGCGCGCAAGCGCGAGGTCTGTACACAGCC 666
Db 6673 CGCAGTGCCTGTGCTGGCGCAGCGCTGGCGCGCAAGCGCGAGGTCTGTACACAGCC 6732
Qy 667 TTCAGGCGCTCGGAGATTTCGGTGGACTACGTGCAGGTCTGT 708
Db 6733 TTCAGGCGCTCGGAGATTTCGGTGGACTACGTGCAGGTGAGT 6774

RESULT 5
US-09-949-016-201775/c
; Sequence 17175, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17175
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-201775

Query Match 37.0%; Score 396.4; DB 4; Length 601;
Best Local Similarity 99.2%; Pred. No. 6.6e-70;
Matches 397; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 309 CTACGAGAGACCCCGGTACTTGGCGCAGTGTGGCGTCCGCCCGCGCCCTGCTGTA 368
Db 601 CTACGAGAGACCCCGGTACTTGGCGCAGTGTGGCGTCCGCCCGCGCCCTGCTGTA 542
Qy 369 CCGCGCGCGCGCTCGCGCTCTCATGTGTGGATGGCAACCGCGCGAGGACCTCTA 428
|
```

```
Db 541 CCCGCGCGCGGCTGGCGTCTCTCATGTGTGGATGGCAACCGCGCGGAGGACCTCTA 482
Qy 429 CATGGTGCACATGTTCCGCGAGAGTCTTCGTGACGAGGACCCCGCCACGTACTGTGGGA 488
|
|
|
Db 481 CATGGTGCACATGTTCCGCGAGAGTCTTCGTGACGAGGACCCCGCCACGTACTGTGGGA 422
Qy 489 CGGCAACTTACCACAGCCCTGGGAACCCCGCGCGCGCGCGTGGCGCGGAGCCTA 548
|
|
|
Db 421 CGGCACTTACCACAGCCCTGGGAACCCCGCGCGCGCGCGTGGCGCGGAGCCTA 362
Qy 549 TCGGAGGTGAGGCGGAGGATCTTGGGCGGCTGGCAGTGGAGCGCTGTGTGAGACTCG 608
|
|
|
Db 361 TCGGAGGTGAGGCGGAGGATCTTGGGCGGCTGGCAGTGGAGCGCTGTGTGAGACTCG 302
Qy 609 CAGTGTGCTGCTGGCGCAGCGCTGGGCGGCGCAAGCGGAGGTCTATGTACAGCCCTT 668
|
|
|
Db 301 YAGTGTGCTGCTGGCGCAGCGCTGGGCGGCGCAAGCGGAGGTCTATGTACAGCCCTT 242
Qy 669 CAAGGCGCTCGGAGATTTCGGTGGACTACGTGCAGGTCTGT 708
|
|
|
Db 241 CAAGGCGCTCGGAGATTTCGGTGGACTACGTGCAGGTGAGT 202

RESULT 6
US-09-949-016-3728
; Sequence 3728, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3728
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3728

Query Match 25.5%; Score 272.8; DB 4; Length 1662;
Best Local Similarity 61.0%; Pred. No. 2.4e-45;
Matches 533; Conservative 0; Mismatches 302; Indels 39; Gaps 4;
Qy 61 CGGAGGGTGTGACCATCGCTTCGCGCTGTCTCATCTCTGGGCTCTCATGACCTGGGCTTAC 120
|
|
|
Db 28 CGTGTGTGGGACACGACCTGTTTCCCTGGCAGTGTGGTGGCATCTTGGCAGCTAT 87
|
|
|
Qy 121 GCGCGCGGGTGGCGCTGGCGCTCGCATCGCTACGGCTCTCTGGCTTCGGCTCTACGG 180
|
|
|
Db 88 GTGACGGGCTACCATGTTTCATCCACACGGAAGACATACCTGTCTTCGGCTGTACGGC 147
|
|
|
Qy 181 GCCTTCCTTTTCAGCGCACCTGGTGGCGCAGAGCTCTTCGGTACCTGGAGCACCGCGG 240
|
|
|
|
Db 148 GCCATCTTGGGCTTGCACCTGTCTCATTCAGAGCCCTTTTTCCTTCCTGGAGCACCGGCG 207
Qy 241 GTGCGCGCGCGCGCGCGGGGCGCTGGATGGACGACCGGCGC---GCAGTGTGGCGCTG 297
|
|
|
|
Db 208 ATGCGACGTGCGCGCCAGCCCTGAAAGTGCCTCCCGCGCGGGGCTCGGTGGCACTG 267
Qy 298 ACATCTTCGCTTACGAGGAGGACCCCGCTGCTGCTGCGGAGTGTGGCTTCGCGCGC 357
|
|
|
|
Db 268 TGCATTGCGCATACACGAGGAGCCCTTGACTTTCGCAAGTGTGCTTCGCTCGGCTCGGCGCAG 327
|
|
|
|
Qy 358 GCCCTGTGTACCGCGCGCGGCTCGCGCTCCTCATGTGTGGATGGCAACCGCGGCC 417
|
```


Db 328 CGCATCTCTCCCTGAC-----CTCAAGTGGTTCATGTTGGTGGTGAACGCCAG 381
Qy 418 GAGGACCTCTACATGGTGCACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCGCCACG 477
Db 382 GAGGACGCTACATGCTGACATCTTCCAGAGGTGCTGGCGGACCGAGGAGCGCGC 441
Qy 478 TACGTGTGGAGCGCAATCAACAGCCCTGGGAAACCGCGCGCGCGCGCGCTGGGC 537
Db 442 TCTTTGTGTGGCGCAGCACTTCCA-----TGAGGCGAGCGAGGTGAG 486
Qy 538 GCCGAGGCTATCGGAGGTGGAGCGGAGGATCTGGCGGCTGGCAGTGGAGCGGTG 597
Db 487 ACAGGAGCGAGCTCGAGAGGCGATGGAACGT-----GTGCGGATGTG 531
Qy 598 GTGAGGACTCGCAGTGTGCTGCTGGCGCAGCGCTGGCGGCGCAAGCGCGAGTCTATG 657
Db 532 GTGCGGCGCAGCACCTTCTGTCATCATGCAAGTGGGAGGCGAGGTGATG 591
Qy 658 TACAGAGCTTCAAGCGCTCGGAGATTCGGTGGACTAGTGGAGTCTGTGACTCGAC 717
Db 592 TACAGGCTTCAAGCGCTCGGCGATTCGGTGGACTACATCCAGGTGTGCGACTCTGAC 651
Qy 718 ACAAGTTGGACCCCATGCTGCTGAGCTCGTGGGTACTGGAAGGACCCCGG 777
Db 652 ACTGTGCTGGATCCAGCTGCAACATCGAGATGCTTCGAGTCTTCGAGGAGGATCCCAA 711
Qy 778 GTAGGGGCTGTGTGGGAGAGCTGGGATCCTTAACCTCTGGACTCTGGGTGAGCTTC 837
Db 712 GTAGGGGAGTGGGGGAGATGTCAAGTCTCAAGATGAGTCAATGATTCCTTC 771
Qy 838 TAAAGCAGCTCGGATCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTC 897
Db 772 CTGAGCAGCTGGCGTACTGGATGCGCTTCAAGTGGAGCGGCGCTGCCAGTCTACTTT 831
Qy 898 CACTGTGTATCTGCAATCAGCGTTCCTAGGTA 931
Db 832 GGCTGTGTGCACTGTATTAGTGGGCGCTTGGGCA 865

RESULT 7

US-08-812-008-31

; Sequence 31, Application US/08812008

; Patent No. 6602693

; GENERAL INFORMATION:

; APPLICANT: McDonald, J. A.

; APPLICANT: Spicer, A. P.

; APPLICANT: Augustine, M. L.

; TITLE OF INVENTION: GENE ENCODING HYALURONAN

; TITLE OF INVENTION: SYNTHASE

; NUMBER OF SEQUENCES: 54

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Schwesman, Lundberg, Woessner & Kluth, P.A.

; STREET: P.O. Box 2938

; CITY: Minneapolis

; STATE: MN

; COUNTRY: U.S.A

; ZIP: 55402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/812,008

; FILING DATE: 05-MAR-1997

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/675,499

; FILING DATE: 03-JUL-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Embretson, Janet E

; REGISTRATION NUMBER: 39,665

; REFERENCE/DOCKET NUMBER: 150.183US1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 612-373-6900

; TELEFAX: 612-339-3061

; TELEX:

; INFORMATION FOR SEQ ID NO: 31:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1665 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: unknown

; TOPOLOGY: unknown

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: Coding Sequence

; LOCATION: 1...1662

; OTHER INFORMATION:

US-08-812-008-31

Query Match 22.1%; Score 236.2; DB 4; Length 1665;

Best Local Similarity 58.4%; Pred. No. 4.5e-38;

Matches 512; Conservative 0; Mismatches 323; Indels 42; Gaps 4;

Qy 61 CGGAGGTGCTGACCATCGCTTCCGCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
Db 28 CGTGTGGTGGGACCGAGTCTGTTTGGCTGTGCTGGAGGATCTCTGGCGGCTAT 87
Qy 121 GCCCGCGGGTGGCGCTGGCTCGATCGCTACGGCTCTCTGGGCTTGGGCTTACGGG 180
Db 88 GTGACAGCTACCAAGTTTATCCACAGAAAGACATACCTGTCTTGGCTCTACGGT 147
Qy 181 GCCTTCTCTTTCAGCGCACTGTGGCGCAGAGCTCTTTCGCTACCTGGAGCACCGGG 240
Db 148 GCCATCTCTGGGTACATCTGCTCATCCAGAGCTGTTTGGCTTCTTGGAGCACGTCGA 207
Qy 241 GTGGCGGGCGGGCGCG-----GGGGCGCTGGATGACGACCGCGCGGAGTGTGGG 294
Db 208 ATGCGCAGGGCAGGGCGGCGCTCAAGCTGCATGCTCCAGAGGTCTGCTTCACTGGCA 267
Qy 295 CTGACCATCTCCGCTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 354
Db 268 CTCTGCAATGCTGCTTACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 327
Qy 355 CGCGCCCTGTGTATACCGCGCGCGGCTGCGCGCTCTCTCATGTGGTGGTGGTGAACCGC 414
Db 328 CAGCGCATGCTCTTCCAAAC-----CTCAAGGTGGTCTATGATGATGATGATGATG 381
Qy 415 CGCGAGGACCTCTACATGCTGATGTTCCGCGAGGTCTTCCGCTGACGAGGACCGCGC 474
Db 382 CAGGAAGATACCTATGTTGGACATCTTCCATGAGGTGCTGGGTGGCACTGAGCAAGCT 441
Qy 475 ACGTACGTGTGGGAGCGCAACTACACAGCCCTGGGAAACCGCGCGCGCGCGCGGTG 534
Db 442 GGCTTCTTGTGGGTAGCAATTTCCA-----TGAGCGGGGTGAAGA 486
Qy 535 GCGCGCGAGGCTTACGCGAGGTGAGCGGAGGATCTTGGCGCGCTGGCAGTGGAGGCG 594
Db 487 GAGACAGAGGCGAGCTGCGAGGAGGCAAGGAGTGGAGGT-----GTGCGAGCT 531
Qy 595 CTGTGTAGGACTCGCAGGTGCTGTGCGCGCAGCGCTGGGCGGCGCAAGCGGAGGTC 654
Db 532 GTGTGTGGGCGAGCAGCTTCTCATGCAATCATGAGAGAGAGAGAGAGAGAGAGAGAG 591
Qy 655 ATGTACACAGCTTCAAGGCGCTCGGAGATTCGGGTGGACTACGTGCAAGTCTGTGACTCG 714
Db 592 ATGTACATGCTTCAAGGCGCTTGGCACTCAGTGGAGTACATCCAGTGTGTGACTCT 651
Qy 715 GACCAAGGTTGGACCCCAATGGCACTGTGAGGCTCTGTCGGGTACTGGAAGGAGACCC 774
Db 652 GACACTGTGTGGACCCAGCTGCGAGGATGCTTGGAGTCTTGGAGTCTTGGAGAGATCCC 711
Qy 775 CGGTAGGGGCTGTGTGGGAGAGTGGGAGTCTTAAACCTCTGGAATCTCTGGGTGAGC 834
Db 712 CAAGTAGGAGGTGTGGAGGAGATGTCCAAATCTCTCAACAAGTATGATTTCATGATCTCC 771

```
Qy 835 TTCCTAAGCAGCTCGCATCTGGGTAGCCTTCAATGTGAGCGGCTTGTGACAGCTAC 894
Db 772 TTCTGTGAGCAGTGTGAGTACTGGATGCTTTCAACGTGGAGCGGCTGCCAGTCTTAC 831
Qy 895 TTCCACTGTGTATCTCTGCATCAGCGGTCTCTAGGTA 931
Db 832 TTTGGCTGTGCAATGTATTAGTGGGCTTTGGCA 868

RESULT 8
US-09-949-016-201780/c
; Sequence 201780, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH TYPE II HYALURONAN, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201780
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; ORGANISM: Human
US-09-949-016-201780
```

```
Query Match 21.5%; Score 230.8; DB 4; Length 601;
Best Local Similarity 98.3%; Pred. No. 4.8e-37;
Matches 232; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 696 CGTCAGGTCTGTGACTCGGACACAAAGTTGGACCCCATGTCAGTCTGGAGCTCGTGG 755
Db 403 CTTACAGGTCTGTGACTCGGACACAAAGTTGGACCCCATGTCAGTCTGGAGCTCGTGG 344
Qy 756 GGTACTGGACGAGACCCCGGGTAGGGCTGTGTGGGACCTGCGGATCCCTTAACCC 815
Db 343 GGTACTGGACGAGACCCCGGGTAGGGCTGTGTGGGAGTGTGGGATCCCTTAACCC 284
Qy 816 TCTGAGCTCTCTGGGTGAGCTTCTTAAGCAGCCTCGCATCTGGGTAGCCTTCAATGTGA 875
Db 283 TCTGAGCTCTCTGGGTGAGCTTCTTAAGCAGCCTCGCATCTGGGTAGCCTTCAATGTGA 224
Qy 876 GCGGGTGTGACAGCTACTTCCACTGTGTATCTGTATCTGATCAGCGGTCTCTAGGTA 931
Db 223 GCGGGTGTGACAGCTACTTCCACTGTGTATCTGTATCTGATCAGCGGTCTCTAGGTA 168
```

```
RESULT 9
US-08-675-499A-1
; Sequence 1, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schweigman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
```

```
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D.
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2947 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-675-499A-1
```

```
Query Match 16.5%; Score 176.2; DB 4; Length 2947;
Best Local Similarity 55.6%; Pred. No. 3.8e-26;
Matches 430; Conservative 0; Mismatches 298; Indels 45; Gaps 3;

Qy 159 CCTGGCTTCGGCTCTTACCGGGGCTTCTTTCAGCGACCTGTGTGGGAGAGCGCTCTT 218
Db 636 CTTCTCATTTGGACTGTACGGTGCCTTTTAGCTCGCATCTCATCATCCAAAGCTCTT 695
Qy 219 CGGTACTCGAGCACCGGGGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 278
Db 696 TGGCTTTTGGAAACACCGGAAATGAAGAGTCCCTTGAACCCCGATTAATTGAACAA 755
Qy 279 CGGGCGCAGTGTGGCGTGCACCTTCCGGCTTACAGAGAGACCCCGCTGACCTGCGCA 338
Db 756 AAGC-----GTAGCACTTGCATCGCTGCTGACCAAGAGACCTGACTACTTACGAA 809
Qy 339 GTGCTGGCTCGCCCGCCCTGTGTGTACCCGGCGCGGGCTGCGCTGCTCATGGT 398
Db 810 ATGTTTGAATCTGTGAAAGGCTGACCTACCCCTG-----GGATTAAAGTCGTGATGGT 863
Qy 399 GGTGGATGGCAACCGCCCGAGGACCTTACATGCTCGACATGTTCCGCGAGGTCTTCGC 458
Db 864 CATGATGGGAACCTCAGACGACGACCTTTACATGATGACATATTCAGCGAAGTTATGG 923
Qy 459 TGACGAGGACCCCGCCACGCTACGTGTGGGACGGCAACTACCACGAGCTCGGAAACCGC 518
Db 924 CAGGACAAATCGGCCACGTACATCTGGAAGAACAACTTTCATGAAAG-----972
Qy 519 GGGCGCGGCGGTGGGGCGCGGAGCTATTCGGGAGGTGGAGCGGAGGATCTCTGGCG 578
Db 973 -----GGACCTGGTGACAGAGAGAGTCCCATGAAGAAAGTTC 1010
Qy 579 GCTGGCAGTGGAGCGCTGTGTGAGGACTCCGAGGTGCTGTGCGTGGCGCAGCGCTGGGG 638
Db 1011 ACAACATGTACCCAAATGGTCTTGTCTAAACAAAGTATTGTGCAATCATGCAAAATGGGG 1070
Qy 639 CGCAAGCGGAGGTATGTACACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTACGT 698
Db 1071 TGGAAAGAGAGAGTGTGTGTACACAGCCTTCAGAGCACTGGGGCGGAAAGCTGGATTGT 1130
Qy 699 GCAGGTCTGTGACTCGGACACAAAGTTGGAGCCCATCGGCACTGTGTGAGCTCGCGGGT 758
Db 1131 ACAGGTGTGACTCAGATATCTATGCTTGGCCCTGCTCATCTCTGTGGAGATGGTGAAGGT 1190
```


QY 61 CGGAGGTGCTGACATGCCCTTCCGCTCTCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
Db 2028 CGTGTGTGGGACACAGCCTCTTTGGCTTGGCACTGTGGTGGCATCTCTGGAGCCTAT 2087
QY 121 GCCCGGGGTGCGCTGGCTCCGATCGCTACCGGCTCTCTGGCTCTTACGGG 180
Db 2088 GTGACGGGCTACGAGTTTCATCCACGGAAGACATACCTGTCTCTGGGCTGTAGGC 2147
QY 181 GCCTTCTCTTTCAGCGCACCTGTGGCGAGAGCTCTTTCGCTTACCTGAGACACCGCGG 240
Db 2148 GCCATCTCTGGGCTGCACTCTCTATTTCAGAGCCTTTTTCCTCTGGAGACACCGCGC 2207
QY 241 GTGG 297
Db 2208 ATGCGAGTGTGG 2267
QY 298 ACCATCTCCGCTTACCAGGAGGACCCCGCTGACCTGGCGGAGTCCCTGGGCTCGCGCGC 357
Db 2268 TGCATTGGCGGATACAGGAGGACCTGACTTCTTGGCAAGTGGCTGGCTGGGCGCAG 2327
QY 358 GCCCTGTGTACCCCGCGCGCGCTGCGCTCTCTCATGTGTGTGTGTGTGTGTGTGTGT 417
Db 2328 CGCATCTCTTCCCTGAC-----CTCAAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2381
QY 418 GAGGACCTTACATGTTGCACTGTTCCGAGAGTCTTCTGCTGACGAGGACCCCGCCACG 477
Db 2382 GAGGACGCTTACATGCTGGACATCTTCCAGAGTGTGTGGCGGACACCGAGCAGCGCGC 2441
QY 478 TACGTGTGGGACGCACTACACAGCCCTGGGAAACCGCGGGGGGGGGGGGGGGGGGGGG 537
Db 2442 TCTTTGTGTGGCGAGCACTTTCAT-----GAGGACGAGGAGGTGAG 2486
QY 538 GCGGAGCCTATCGGAGGTGGAGCGAGAGATCTCTGGCGGCTGGGAGTGGAGCGGCTG 597
Db 2487 ACGGAGCCAGCTTCCAGGAGGAGTGGACCGT-----GTGCGGGATGTG 2531
QY 598 GTGAGGACTCGAGGT 657
Db 2532 GTGGGGGGCGACACCTTCTCGTGTATCATGACAGAGTGGGAGGACGAGCGAGGTCATG 2591
QY 658 TACAGAGCTTCAAGGGCTCGGAGATTCTGGTGGACTACGTGCAAGTCTGTG 709
Db 2592 TACAGGCTTCAAGGGCTCGGAGATTCTGGTGGACTACGTGCAAGTAAAGG 2643

RESULT 12

US-09-949-016-737
; Sequence 737, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 737
; LENGTH: 3003
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-737

Query Match 13.9%; Score 148.8; DB 4; Length 3003;
Best Local Similarity 53.5%; Pred. No. 1e-20;

Matches 411; Conservative 0; Mismatches 312; Indels 45; Gaps 3;
QY 164 CTTTGGCCTTACCGGGCCTTCTCTTTCAGGACACCTGGTGGCGCAGAGCTCTTCGGGT 223
Db 669 CTTTGGACTGTATGGTGGCCTTTTGGCATCACACCTCATCATCAAAGCCTGTTCGCT 728
QY 224 ACCTGGAGACCGCGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 283
Db 729 TTTTGGAGCACGGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 782
QY 284 GCAGTGTGGCGCTGACCATCTCCGCTTACAGGAGGACCCCGGTACTCTGGCGCAGTCC 343
Db 783 AAACAGTTGCCCTTTGTCATCTGCTGCTTATCAAGAGATCCAGACTACTTAAAGGAAAT 842
QY 344 TGGGTCCG 403
Db 843 TGCATCTGTGAAAGGCTAACCTACC-----CTGGGATTAAGTTGTATGGTTCATAG 896
QY 404 ATGGCAACCGCGCGAGGACCTCTACATGTTCGACATGTTCCGCGAGGTCTTCGCTGACG 463
Db 897 ATGGGAACCTCAGAAGATGACCTTTTACATGATGACATCTTTCAGTGAAGTCAATGGCAG 956
QY 464 AGGACCCCGCAGCTAGTGTGGAGCGGCAACTACCAACAGCCCTGGGAAACCCCGCGCGG 523
Db 957 ACAAATCAGCCACTTATCTGGAAGAACTTCCACAGAAAG----- 1000
QY 524 CGGGCGGTGGGCGCGGAGCCTATCGGAGGTGGAGGCGGAGGATCTCTGGGCGGCTGG 583
Db 1001 -----GGTCCCGGTGAGACAGATGATGATCAATAAAGAAAGCTTCGCAAC 1043
QY 584 CAGTGGAGCGCTGGTGGAGCTCGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGCGCGCA 643
Db 1044 ACGTAACGAATTTGCTTCTGTCACAAAGATATCTCATCATGCAAAATGGGGTGGAA 1103
QY 644 AGGCGGAGTCTATGTACACAGCCTTCAAGGCGCTCGGAGATTTCGTTGAGACTACGTGCG 703
Db 1104 AAGAGAAAGTCATGTACACAGCCTTCAGAGCCTTCAGAGCAGTGGAGTGTGATTTATG 1163
QY 704 TCTGTGACTCGGACACAGGTTGGACCCCATGGACTGTCTGGAGCTCGTGGGGTACTGG 763
Db 1164 TTTGTGATTGAGACACTATGCTTTGACCCAGCCTCATCTGTGGAGATGTTAAAGTTT 1223
QY 764 ACAGAGACCCCGGTGAGGGCTGTGTTGGTGGGACGTCGCGATCCTTAACCTCTGACT 823
Db 1224 AAGAATCCCATGTTGGAGGTGTTGGGGAGATGTCCAGATTTTAAACAAGTACGATT 1283
QY 824 CTTGGGTGAGCTTCTTAAGCAGCCTGCGATATCTGGGTAGCCTTCAATGTGGAGCGGCT 883
Db 1284 CTTGGATCTCATTTCTCAGCAGTGTAAAGATATTGGATGGCTTTTAAATATAGAAAGG 1343
QY 884 GTGAGACTACTTCCACTGTGTATCTGCAATCAGCGGTTCTCTAGGTA 931
Db 1344 GTCAGTCTTATTTGGGTGTGTTTCAAGTGCATTAGTGAGCCTCTGGGAA 1391

RESULT 13

US-09-949-016-132199
; Sequence 132199, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012

```

RESULT 15
; US-09-902-540-938
; Sequence 938, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hankie, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Gen
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10

```

; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 938
; LENGTH: 10210
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-938

Query Match 7.7%; Score 82.8; DB 4; Length 10210;
Best Local Similarity 44.8%; Pred. No. 1.5e-07;
Matches 357; Conservative 0; Mismatches 437; Indels 2; Gaps 1;

QY	3	GAGACAGCAGGACGCGCCCAAGCCACTCCCTGACGCCGCGCTCTCGGGCTGGCCCG	62
DB	4198	GAGGACGCGCCACCGCGAGGTGCTCTACCGCCGCGGGGACCTCTTCGACGACCCG	4257
QY	63	GAGGTGCTGACCATCGCC--TTGCCCTGTCTCATCTGGGCTCATGACCTGGGCTTAC	120
DB	4258	AAGGACGCGCATCGACGCGCTGTTCCGCTGGCGCCTGCGTCCCGGATGCGTCCGTC	4317
QY	121	GGCGCGGGGTGCGCTCGCTCCGATCGCTACGGCTCTCTGGCTTCGGCTCTACGGG	180
DB	4318	ATCGACCGGGCGGTGGCGGGGTGCACGCCCTGGGACGTCATGGCGACCTGCTGGACGTG	4377
QY	181	GCTTCTCTTCAGCGCACTCGTGGCGCAGAGCCTCTTCGGGTACTCTGAGCACCGCGG	240
DB	4378	TACGCTCGGGCGCGAGCGCGGGTGGCGCTCGCGCGCGAGCTGCTGTGGCCGCC	4437
QY	241	GTGGCGGGCGCGCGGGGCGCTGGATGACGACACCGCGCGCAGTGTGGCGCTGACC	300
DB	4438	GGCTCGGTGGCTCTCAGCTCGCTGGCGGATGAAGACGCGGCTGGACCTGACCCAGCGC	4497
QY	301	ATCTCGGCTTACAGAGGACCCCGGTAACCTGGCGCAGTGGCTGGCGCTCGCCCGGCC	360
DB	4498	GGCGCGGAGCGCGCCGCGAGACCTCACCGCGCTCCAGGCACCTGGTGAACGGGCTGCGT	4557
QY	361	CTGCTGTACCGCGCGCGCTCGCTCGCTCTCATGTGGTGGATGGCAACCGCGCCGAG	420
DB	4558	CAGCGGGCGGACTCACCGGCTGCTGGAGGGCTGGAGCAGTGGGTGCGCGGCTCGAG	4617
QY	421	GACCTCTATATGTGTCGATGTTCCGCGAGGTCCTTCGTGACGAGGACCCCGCCACGTAC	480
DB	4618	GACGCCGACGAGCGCGGTGCTGCGGCTGGAGCTGGCCGAGCTCGCGCTGACGCGGC	4677
QY	481	GTGTGGAGCGCACTACACAGCCCTGGGNAACCGCGCGCGGGCGCGCTGGGGGCC	540
DB	4678	CGGTGGATGTGGCGCGGAGCACTGGAGGAAGTGGCGCGCGCTGGCGGCGG	4737
QY	541	GGAGCTATCGGAGGTGGAGCGGAGGATCTTGGGCGGCTGGCAGTGGAGCGCTGGTG	600
DB	4738	GGTTACGGGACGCGCTGGAGCCCTCGAGCGCTGCTGAAGGACGCCCGCCCGCGCGC	4797
QY	601	AGGACTCGAGTGTGCTGGTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCTATGTAC	660
DB	4798	GCCGAGGTGCAAGTGGCGCGCGAGTTGGCTTCGGCGCGGACGCGGAGGTGCTGCTC	4857
QY	661	ACAGCTTCAAGCGCTCGAGATTCCGTGGACTACGTGACAGTCTGTGACTCGGACACA	720
DB	4858	ATGGCGCGCGCGGGGTTGAGTTCGCGCGGACAGCTTCCGGAGGGCCCTGAAGCGCGG	4917
QY	721	AGTTTGGACCCATCGCACTGCTGGAGCTCTGTGGGTACTGGACGAGGACCCCGGGTA	780
DB	4918	AAGNCGCGCTCGGTGGAGCCGACGCTGACCGCGCCCTGCGAGTGGCCACCTGTAC	4977
QY	781	GGGCTGTTGGGGG	796
DB	4978	CGCGCTCGGCGAGG	4993

Search completed: March 13, 2005, 08:12:15
Job time : 217.96 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:08:52 ; Search time 4414.94 Seconds
(without alignments)
9233.835 Million cell updates/sec

Title: US-10-672-399-3
Perfect score: 1071
Sequence: 1 atgagacagcaggacgcgc.....ccggcaccatgcgtggatga 1071

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_hic:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gsa1:
9: gb_gsa2:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	916.8	85.6	2072	3	CR602106	CR602106 full-length	
2	915.2	85.5	1728	9	AY421569	AY421569 Homo sapi	
3	822.4	76.8	1058	4	BM543754	BM543754 AGENCOURT	
4	795.6	74.3	965	1	AL545447	AL545447 AL545447	
5	679.8	63.5	1746	9	AY421571	AY421571 Mus muscu	
C	6	490.4	45.8	896	5	BM371636	BM371636 BX371636
7	479.4	44.8	3542	3	AK053726	AK053726 Mus muscu	
8	357.4	33.4	650	2	BB665418	BB665418 BB665418	
9	305.8	28.6	1084	4	BM544718	BM544718 AGENCOURT	
10	291.8	27.2	551	4	BG707429	BG707429 602672848	
11	291.8	27.2	551	4	BG707608	BG707608 602670672	
12	268.2	25.0	508	9	CE377866	CE377866 tigr-g88-	
13	258.2	24.2	610	5	BX327795	BX327795 BX327795	
14	256.2	23.9	894	7	CF239758	CF239758 AGENCOURT	
15	255.8	23.9	782	4	B1753116	B1753116 603025961	
C	16	253.6	23.7	668	8	BH057613	BH057613 RPC1-24-2
17	237.8	22.2	774	5	BP677403	BP677403 BP677403	
18	236.2	22.1	4151	3	AK028582	AK028582 Mus muscu	
19	223.8	20.9	895	5	BU911059	BU911059 AGENCOURT	
20	209	19.5	747	6	CD466132	CD466132 Leukon2.2	
21	206.8	19.3	604	4	BG016146	BG016146 deS7d05.y	
22	202.6	18.9	944	6	CA487249	CA487249 AGENCOURT	
C	23	194.2	18.1	495	7	CN664556	CN664556 A0814F02-
24	186.2	17.4	660	5	BP674137	BP674137 BP674137	

ALIGNMENTS

RESULT 1
CR602106 2072 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CSODI027YJ02 of Placenta Cot 25-normalized
DEFINITION full-length cDNA clone CSODI027YJ02 of Placenta Cot 25-normalized
ACCESSION CR602106
VERSION CR602106.1 GI:50482913
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li W.B., Gruber C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 2072)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..2072
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI027YJ02"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 85.6%; Score 916.8; DB 3; Length 2072;
Best Local Similarity 99.8%; Pred. No. 2.6e-174;
Matches 918; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 10 CAGGACGGCCCAAGCCCACTCTCTGACGCCCGCGCTCTCGGCTCGCCGCGAGGTG 69
=====|=====|=====|=====|=====|=====|=====|=====|=====|=====|
DB 58 CAGGACGGCCCAAGCCCACTCTCTGACGCCCGCGCTCTCGGCTCGCCGCGAGGTG 117

Qy 70 CTGACCATCGCTTGGCCCTGCTCATCTGGCCCTCATGACTGGGCTTACGCGCCGGG 129
 Db 118 CTGACCATCGCTTGGCCCTGCTCATCTGGCCCTCATGACTGGGCTTACGCGCCGGG 177
 Qy 130 GTGCCGCTGGGCTCCGATCGCTACGGCTCTTGGGCTTCCGGCTTACGCGGCTTCTT 189
 Db 178 GTGCCGCTGGGCTCCGATCGCTACGGCTCTTGGGCTTCCGGCTTACGCGGCTTCTT 237
 Qy 190 TCAGCGCACTGGTGGCGAGAGCTTCTTCGCTTACCTGGAGCAACCGCGGCTGGCGCG 249
 Db 238 TCAGCGCACTGGTGGCGAGAGCTTCTTCGCTTACCTGGAGCAACCGCGGCTGGCGCG 297
 Qy 250 GCGCGCGCGGCGCTGGATGCAAGCAACCGCGGCTGGGCTTACCATCTCCGCC 309
 Db 298 GCGCGCGCGGCGCTGGATGCAAGCAACCGCGGCTGGGCTTACCATCTCCGCC 357
 Qy 310 TACAGAGAGACCCCGCTACCTGGCGCAGTGGCTTGGCGCTGGCGGCTTCTGTAC 369
 Db 358 TACAGAGAGACCCCGCTACCTGGCGCAGTGGCTTGGCGCTGGCGGCTTCTGTAC 417
 Qy 370 CCGCGCGCGGCTGGCTTCTCATGCTGGTGGATGCAACCGCGGCTGGAGACTTCAC 429
 Db 418 CCGCGCGCGGCTGGCTTCTCATGCTGGTGGATGCAACCGCGGCTGGAGACTTCAC 477
 Qy 430 ATGCTCGACATGTTCCGCGAGGCTTCTGCTGACGAGACCCCGCACTGCTGGGAC 489
 Db 478 ATGCTCGACATGTTCCGCGAGGCTTCTGCTGACGAGACCCCGCACTGCTGGGAC 537
 Qy 490 GGCAACTACCAACAGCCCTGGGAAACCCCGCGGCGGCGGCTGGGCGCGGAGCTTAT 549
 Db 538 GGCAACTACCAACAGCCCTGGGAAACCCCGCGGCGGCGGCTGGGCGCGGAGCTTAT 597
 Qy 550 CGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCACTGGAGGCGCTGGTGGAGACTCGC 609
 Db 598 CGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCACTGGAGGCGCTGGTGGAGACTCGC 657
 Qy 610 AGGTGCTGCTGGTGGCGGAGGCTTGGCGGCGGCGGCGGAGGCTGATGACAGGCTTC 669
 Db 658 AGGTGCTGCTGGTGGCGGAGGCTTGGCGGCGGCGGCGGAGGCTGATGACAGGCTTC 717
 Qy 670 AAGCGCTCGGAGATTCGGTGGAGCTTACGTGAGGCTTGTGACTCGGACACAGGTTGGAC 729
 Db 718 AAGCGCTCGGAGATTCGGTGGAGCTTACGTGAGGCTTGTGACTCGGACACAGGTTGGAC 777
 Qy 730 CCATGAGCACTGCTGGAGCTCGTGGGTA CTGACGAGGACCCCGGCTAGGGGCTGTT 789
 Db 778 CCATGAGCACTGCTGGAGCTCGTGGGTA CTGACGAGGACCCCGGCTAGGGGCTGTT 837
 Qy 790 GGTGGGAGCTGGGATCTTAACCTCTGACTTCTGGGCTAGGCTTCTTAAGCAGCTG 849
 Db 838 GGTGGGAGCTGGGATCTTAACCTCTGACTTCTGGGCTAGGCTTCTTAAGCAGCTG 897
 Qy 850 CGATACCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCACTGTGTATCC 909
 Db 898 CGATACCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCACTGTGTATCC 957
 Qy 910 TGCATCAGCGCTTCTTAGG 929
 Db 958 TGCATCAGCGCTTCTTAGG 977

RESULT 2
 AY421569
 LOCUS 1728 bp DNA linear GSS 17-DEC-2003
 DEFINITION Homo sapiens HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY421569
 VERSION AY421569.1 GI:39748428
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 1728)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)
 14671302
 2 (bases 1 to 1728)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source
 1..1728
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 gene
 <1..>1728
 /gene="HAS1"
 /locus_tag="HOM7601"
 ORIGIN
 Query Match 85.5%; Score 915.2; DB 9; Length 1728;
 Best Local Similarity 99.7%; Pred. No. 5.4e-174;
 Matches 917; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 Qy 10 CAGGACCGCGCCCAAGCCCACTCTCTGACGCGCGCTCTCGGCTTGGCGGCTTCTT 69
 Db 1 CAGGACCGCGCCCAAGCCCACTCTCTGACGCTCTCTGCGCTTGGCGGCTTCTT 60
 Qy 70 CTGACCATCGCTTGGCCCTGCTCATCTGGGCTTCTGAGCTTGGGCTTACGCGCCGGG 129
 Db 61 CTGACCATCGCTTGGCCCTGCTCATCTGGGCTTCTGAGCTTGGGCTTACGCGCCGGG 120
 Qy 130 GTGCCGCTGGGCTCCGATCGCTACGCGCTCTCTGCGCTTGGCGCTTCTTCTT 189
 Db 121 GTGCCGCTGGGCTCCGATCGCTACGCGCTCTCTGCGCTTGGCGCTTCTTCTT 180
 Qy 190 TCAGCGCACTGGTGGCGGAGAGCTTCTTCTGCTTACCTGGAGCAACCGCGGCTGGCG 249
 Db 181 TCAGCGCACTGGTGGCGGAGAGCTTCTTCTGCTTACCTGGAGCAACCGCGGCTGGCG 240
 Qy 250 GCGCGCGCGGCGCTGGATGCAAGCAACCGCGGCTGGGCTTGGGCTTACCATCTCCGCC 309
 Db 241 GCGCGCGCGGCGCTGGATGCAAGCAACCGCGGCTGGGCTTGGGCTTACCATCTCCGCC 300
 Qy 310 TACGAGAGGACCCCGGCTACCTGCGCAGTGGCTTGGCTTGGCTTGGCTTGGCTTAC 369
 Db 301 TACGAGAGGACCCCGGCTACCTGCGCAGTGGCTTGGCTTGGCTTGGCTTGGCTTAC 360
 Qy 370 CCGCGCGCGGCTGGCGCTCTCTCATGGTGGATGGCAACCGCGGCTGGAGAGCTTCTTAC 429
 Db 361 CCGCGCGCGGCTGGCGCTCTCTCATGGTGGATGGCAACCGCGGCTGGAGAGCTTCTTAC 420
 Qy 430 ATGCTCGACATGTTCCGCGAGGCTTCTGCTGACGAGGACCCCGGCTTGGGCTGGGAC 489
 Db 421 ATGCTCGACATGTTCCGCGAGGCTTCTGCTGACGAGGACCCCGGCTTGGGCTGGGAC 480
 Qy 490 GGCAACTTACCAACAGGCTTGGGAAACCCCGCGGCTGGGCTTGGGCTTGGGCTTAT 549
 Db 481 GGCAACTTACCAACAGGCTTGGGAAACCCCGCGGCTGGGCTTGGGCTTGGGCTTAT 540
 Qy 550 CCGGAGGTGGAGCGGAGGATCTTGGGCGGCTTGGCACTGGAGGCTTGGTGGAGACTCGC 609
 Db 541 CCGGAGGTGGAGCGGAGGATCTTGGGCGGCTTGGCACTGGAGGCTTGGTGGAGACTCGC 600


```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE       Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL     Full-length cDNA libraries and normalization
COMMENT     Unpublished (2001)
            On Feb 15, 2001 this sequence version replaced gi:31267282.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            9502.r
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?r=CS0D1027DE01QPI&c=9502.r.
FEATURES    Location/Qualifiers
            1..965
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CS0D1027J02"
            /tissue_type="PLACENTA COT 25-NORMALIZED"
            /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
            /note="1st strand cDNA was primed with a NotI-oligo(dT)
            primer. Five prime end enriched, double-strand cDNA was
            digested with Not I and EcoR V sites of the pCMVSPORT 6 vector.
            Library was normalized."
ORIGIN
Query Match      74.3%; Score 795.6; DB 1; Length 965;
Best Local Similarity 97.4%; Pred. No. 6.4e-150;
Matches 851; Conservative 9; Mismatches 8; Indels 6; Gaps 5;
Qy 10 CAGACGCGCCCAAGCCACCTCTCGACGCGCGCGCTCTCGGCTGCGCGGAGGGTG 69
Db 58 CAGACGCGCGCCCAAGCCACCTCTCGACGCGCGCGCTCTCGGCTGCGCGGAGGGTG 116
Qy 70 CTGACCATCGCTTCGCGCTCATCTCGGCTCATGACCTGGCTACGCGCGCGGG 129
Db 117 CTGACCATCGCTTCGCGCTCATCTCGGCTCATGACCTGGCTACGCGCGCGGG 176
Qy 130 GTGCGCTGCGCTCGATCGCTACGGCTCTCGGCTTCGGCTTACGGGCGCTTCCTT 189
Db 177 GTGCGCTGCGCTCGATCGCTACGGCTCTCGGCTTCGGCTTACGGGCGCTTCCTT 236
Qy 190 TCAGCGCACTGTGGCGCAGAGCTCTTCGCTACCTTGAGCACCAGCGCGGTGGCGCG 249
Db 237 TCAGCGCACTGTGGCGCAGAGCTCTTCGCTACCTTGAGCACCAGCGCGGTGGCGCG 296
Qy 250 GCGCGCGCGCGCGCTGGATGACGCCACCGCGCGCAGTGGCGCTGACCATTCGCGC 309
Db 297 GCGCGCGCGCGCGCTGGATGACGCCACCGCGCGCAGTGGCGCTGACCATTCGCGC 356
Qy 310 TACCAGAGGACCCCGCTACTCGCGCAGTGCCTGGGCTCGCGCGCGCTGCTGTAC 369
Db 357 TACCAGAGGACCCCGCTACTCGCGCAGTGCCTGGGCTCGCGCGCGCTGCTGTAC 416
Qy 370 CCGCGCGCGCGCTGCGCGTCTCATGCTGTGATGCGCAACCGCGCGGAGCACTCTAC 429
Db 417 CCGCGCGCGCGCTGCGCGTCTCATGCTGTGATGCGCAACCGCGCGGAGCACTCTAC 476
Qy 430 ATGCTGCACATGTTCCGAGAGTCTTCGCTACGAGGACCCCGCACAGCTGTGGAC 489
Db 477 ATGCTGCACATGTTCCGAGAGTCTTCGCTACGAGGACCCCGCACAGCTGTGGAC 536
Qy 490 GGCAACTACCAACGAGCCCTGGGAACCCCGCGCGCGCGGTGGCGCGCGAGCCTAT 549

```

```

Db 537 GGCAACTACCAACGAGCCCTGGGAACCCCGCGCGCGGTGGCGCGCGAGCCTAT 596
Qy 550 CGGAGGTGGAGGCGGAGGATCTCTGGGCGGCTGCGAGTGGAGGCGCTGTTGAGGACTCGC 609
Db 597 CGGAGGTGGAGGCGGAGGATCTCTGGGCGGCTGCGAGTGGAGGCGCTGTTGAGGACTCGC 656
Qy 610 AGGTGCGTGTGCTGGCGCAGCGCTGGGCGGCAAGCGCGAGGTGTCATGTACACAGCCCTTC 669
Db 657 AGGTGCGTGTGCTGGCGCAGCGCTGGGCGGCAAGCGCGAGGTGTCATGTACACAGCCCTTC 716
Qy 670 AAGCGCTCGGAGATTCGGTGGACTACGTGACAGTCTGTGTGACTCGGACACAAAGTTGGAC 729
Db 717 AAGCGCTCGGAGATTCGGTGGACTACGTGACAGTCTGTGTGACTCGGACACAAAGTTGGAC 776
Qy 730 CCATGCGACTGCTGGAGCTGCTGCGGTTACTGACAGGACCCCGGTTAGGCGCTTT 789
Db 777 CCATGCGACTGCTGGAGCTGCTGCGGTTACTGACAGGACCCCGGTTAGGCGCTTT 836
Qy 790 GGTGGGACGTGGCGATCTTAACCTCTGACCTCTGGGTGAGTCTTCTAAGCAGCCTG 849
Db 837 GGTGGGACGTGGCGATCTTAACCTCTGACCTCTGGGTGAGTCTTCTAAGCAGCCTG 892
Qy 850 CGATACTGGGTAGCTTCAATGTGGAGCGGCTT 883
Db 893 CGATACTGGGTAG-CTTAAATGTGGAGCGGTTK 925

RESULT 5
AY421571 1746 bp DNA linear GSS 17-DEC-2003
LOCUS Mus musculus HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY421571
VERSION AY421571.1 GI:39748430
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1746)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,
Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1746)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,
Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES Location/Qualifiers
            source 1..1746
            /organism="Mus musculus"
            /mol_type="genomic DNA"
            /db_xref="taxon:10090"
            <1..>1746
            /gene="HAS1"
            /locus_tag="HCM7601"
ORIGIN
Query Match      63.5%; Score 679.8; DB 9; Length 1746;
Best Local Similarity 84.6%; Pred. No. 1.3e-126;
Matches 794; Conservative 0; Mismatches 127; Indels 18; Gaps 2;
Qy 10 CAGGACGCGCCCAAGCCCACTCTCTGACGCGCGCGCTGCTCCGCGCTGCGCGGAGGTG 69

```


Db 419 CCTCTGAGCTCCTGGGTGAGCTTCTTAAGCAGCCTCGCATCTAGGTAGCCTTCAATGTG 360

Qy 874 GAGCGGCTCTCAGAGCTACTTCCACGTGTATCTTCATCAGCGGTCTCTAGG 929
|||||

Db 359 GAGCGGCTTGTGAGAGCTACTTCCACGTGTATCTTCATCAGCGGTCTCTAGG 304
|||||

RESULT 7

AK053726

LOCUS

DEFINITION AK053726 3542 bp mRNA linear HTC 03-APR-2004

AK053726 Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:EI30302P16 product:hyaluronan synthase1, full insert sequence.

ACCESSION AK053726

VERSION AK053726.1 GI:26343670

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 Carninci, P. and Hayashizaki, Y.

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

PUBMED 10349636

REFERENCE 2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE 3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watanuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuina, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watanuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE 4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

MEDLINE 11076861

REFERENCE 5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

MEDLINE 11076861

REFERENCE 6 (bases 1 to 3542)

AUTHORS Fukuda, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Adachi, J., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayate, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kaga, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saigo, R., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, I., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

AUTHORS Muramatsu, M. and Hayashizaki, Y.

TITLE Direct Submission

JOURNAL Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

Location/Qualifiers

1. 3542

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:EI30302P16"

/db_xref="taxon:10090"

/clone="EI30302P16"

/tissue type="eyeball"

/clone lib="RIKEN full-length enriched mouse cDNA library"

/dev stage="0 day neonate"

61. 456

/note="unnamed protein product; hyaluronan synthase1 (MDJ|MG1:106590, GB|NM_008215, evidence: BLASTN, 99%, match=762) putative"

/codon start=1

/protein id="BAC35492.1"

/db_xref="GI:26343671"

/translation="MRQDMKPKSEARCCSGLARRALTIIFALLILGLMTWAYAGVP LADRYGLLAFLGILYAFSLVAHLVAQSLFAYLEHRRVAAARRSLAKPLDAATRSVA LTI SAYQEDPAYLRQCLTSARALLYPHT"

ORIGIN

Query Match 44.8%; Score 479.4; DB 3; Length 3542;

Best Local Similarity 81.6%; Pred. No. 2.7e-86;

Matches 598; Conservative 0; Mismatches 116; Indels 19; Gaps 3;

Qy 10 CAGGACGCGCCCAAGCCCACTCTCTCAGCGCGCGCTGCTCCGCGCTCGCGCGCGGAGGTG 69
|||||

Db 67 CAGGACATGCAAGAGCCCTCAGAGGAGCGGTTGCTGCTCTGCGCTGGCAGCGAGCA 126
|||||

Qy 70 CTGACCATCGCCTTCGCGCTGCTCATCTCGGCGCTCATGACCTGCGGCTACGCGCGCGG 129
|||||

Db 127 CTCACGATCATCTTGGCTGCTCATCTCGGCGCTCATGACCTGCGGCTACGCGCGCAGC 186
|||||

Qy 130 GTGCGCTGCGCTCCGATCGCTACGCGCTCTGCGCTTCGCGCTTACGCGCGCTTCTT 189
|||||

Db 187 GTTCTCTGCTTCAGATCGCTATGGACTCTGCGCTTTCGCGCTTATGGCGCATCTCTC 246
|||||

Qy 190 TCAGCGCACCCTGTGGCGCAGAGCCTCTTCGCGTACCTGAGACACCGCGCGGTGCGCGG 249
|||||

Db 247 AGGCGACACCTAGTGGCAGAGAGCCCTTTCGCTTACCTGAGGACCGAGGTTGCGAGC 306
|||||

Qy 250 GCGGCGCG-----GGGCGCGCTGATGACGACCGCGCGCGAGTGTGGCG 294
|||||

Db 307 GCTGCGCGCGCTCTCTTGGCGAAGGGCGCTGATGCGGCGCTACGCGCGCGCGCA 366
|||||

Qy 295 CTGACCATCTCCGCTTACGAGGAGGACCGCGCTTACCTGCGCGCGCTTCTGCGCGCTCGCC 354
|||||

Db 367 CTACCATCTCAGCCTTACCAAGAGGATCCGCGCTTACCTGCGCGCGCTTGTGACCTCCGCG 426
|||||

Qy 355 CGGCGCTCTGTATCCCGCGCGCGCTGCGCGTCTCATGTGTGTGTGTCGCAACCGC 414
|||||

Db 427 CGGCGCTTGTGTATCCCGCAGACAGAGTAACGCGTGTCTATGTGTGTGAGCGCAACCGC 486
|||||

Qy 415 GCGGAGAGCCTCTACATGTGTGAGATGTTCCGCGAGGTTTCGCTGACGAGGACCGCGCC 474
|||||

Db 487 GCTGAGGATCTGTACATGTGTGAGATGTTCCGAGAGTCTTTCGCGCATGAGGACCGCGCC 546
|||||


```
Db      607 TCGGTGAAGGCGCTACCGGGAGGTGAACGCGAAGACCCCGGGCG 650

RESULT 9
LOCUS   BM544718                      1084 bp      mRNA      linear      EST 20-FEB-2002
DEFINITION   AGENCOURT_6494603 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5727446
5', mRNA sequence.
ACCESSION   BM544718
VERSION     BM544718.1 GI:18776197
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1084)
AUTHORS    NIH-MGC http://mgi.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Invitrogen
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM12721 row: i column: 15
            High quality sequence start: 5
            High quality sequence stop: 689.
            Location/Qualifiers
              1..1084
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5727446"
                /tissue_type="hippocampus"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_124"
                /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcorV
                (destroyed); Site 2: NotI; RNA source male hippocampus,
                age 27. Library is oligo-dT primed and directionally
                cloned (EcorV site is destroyed upon cloning). Average
                insert size 1.4 kb, insert size range 0.9-4 kb. Library is
                normalized and enriched for full-length clones and was
                constructed by C. Gruber (Invitrogen). Research Genetics
                tracking code 012."

FEATURES             source
    source
    1..551
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      /clone="IMAGE:4795643"
      /tissue_type="hypothalamus"
      /lab_host="DH10B"
      /clone_lib="NIH_MGC_96"
      /note="Organ: brain; Vector: pBluescriptR (modified
      pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
      (gtagag); Oligo-dT primed using primer
      5'-TTTTTTTTTTTTTNN-3', size-selected for average
      insert size 2.3 kb and normalized to ROT 5. This is a
      primary library enriched for full-length clones and
      constructed using the Cap-trapper method (Carninci, in
      preparation). Library constructed by M. Brownstein
      (NIH/NHGRI, National Institutes of Health). Note: this is
      a NIH_MGC Library."

ORIGIN
Query Match      28.6%; Score 305.8; DB 4; Length 1084;
Best Local Similarity 99.4%; Pred. No. 2.4e-51;
Matches 307; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      621  CGTGGCGAGCGCTGGGGCGCAAGCGGAGGTGATGTACACAGCCTTCAGGCGCTCGG 680
        |||||
Db       20  CGTGGCGAGCGCTGGGGCGGTAAGCGCGAGGTGATGTACACAGCCTTCAGGCGCTCGG 79
        |||||
QY      681  AGATTTCGGTGACTACGTGACAGGTCTGTGACTCGGACACAAGGTTGGACCCCATGGCACT 740
        |||||
Db       80  AGATTTCGGTGACTACGTGACAGGTCTGTGACTCGGACACAAGGTTGGACCCCATGGCACT 139
        |||||
QY      741  GCTGGAGCTCGTGGGGTACTGGACGAGGACCCCGGGTAGGGCTGTGTGGTGGGACGT 800
        |||||
Db       140  GCTGGAGCTCGTGGGGTACTGGACGAGGACCCCGGGTAGGGCTGTGTGGTGGGACGT 199
        |||||
QY      801  GCGATCCTTAACCTCTGGACTCTGGGTGAGTCTCTAGACAGCTCGGATCTGGGT 860
        |||||
Db       200  GCGATCCTTAACCTCTGGACTCTGGGTGAGTCTCTAGACAGCTCGGATCTGGGT 259
        |||||
QY      861  AGCCTTCAATGTGAGCGGGCTTGTCAAGAGTACTTCCACTGTGTATCTCTGATCAGCGG 920
        |||||

Db      260 AGCCTTCAATGTGAGCGGGCTTGTCAAGAGTACTTCCACTGTGTATCTCTGATCAGCGG 319

QY      921  TTCTCTAGG 929
        |||||
Db       320 TCCTCTAGG 328

RESULT 10
LOCUS   BG707429                      551 bp      mRNA      linear      EST 07-MAY-2001
DEFINITION   602672848F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4795643 5',
mRNA sequence.
ACCESSION   BG707429
VERSION     BG707429.1 GI:13983769
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 551)
AUTHORS    NIH-MGC http://mgi.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1AM10678 row: p column: 12
            High quality sequence stop: 551.
            Location/Qualifiers
              1..551
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4795643"
                /tissue_type="hypothalamus"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_96"
                /note="Organ: brain; Vector: pBluescriptR (modified
                pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
                (gtagag); Oligo-dT primed using primer
                5'-TTTTTTTTTTTTTNN-3', size-selected for average
                insert size 2.3 kb and normalized to ROT 5. This is a
                primary library enriched for full-length clones and
                constructed using the Cap-trapper method (Carninci, in
                preparation). Library constructed by M. Brownstein
                (NIH/NHGRI, National Institutes of Health). Note: this is
                a NIH_MGC Library."

ORIGIN
Query Match      27.2%; Score 291.8; DB 4; Length 551;
Best Local Similarity 99.3%; Pred. No. 1.6e-48;
Matches 293; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      635  GGGCGGCAAGCGGAGGTGATGTACACAGCCTTCAAGGCGCTCGGAGATTCGGTGGACT 694
        |||||
Db       6   GGGCGGCAAGCGGAGGTGATGTACACAGCCTTCAAGGCGCTCGGAGATTCGGTGGACT 65
        |||||
QY      695  ACGTGACAGGTCTGTGACTCGGACACAAGGTTGGACCCCATGGCACTCTGAGCTCGTGC 754
        |||||
Db       66  ACGTGCAGGTCTGTGACTCGGACACAAGGTTGGACCCCATGGCACTCTGAGCTCGTGC 125
        |||||
QY      755  GGGTACGACGAGGACCCCGGGTAGGGCTGTGTGGTGGGACGTCGGATCCTTAACC 814
        |||||
Db       126  GGGTACGACGAGGACCCCGGGTAGGGCTGTGTGGTGGGACGTCGGATCCTTAACC 185
        |||||
QY      815  CTCTGGACTCCTGGGTGAGCTTCTCAAGACGAGCCTGCGATATCTGGGTAGCCTTCAATGTGG 874
        |||||
```


186 CTCTGGACTCTCGGTACGCTTCCTTAAGCAGCCTCGGATCTGGGTAGCCTTCAATGTGG 245
875 AGCGGGCTTGTACAGACTTCTTCCACTGTGTATCTCGCATCAGCGGTTCTCTAGG 929
246 AGCGGGCTTGTACAGACTTCTTCCACTGTGTATCTCGCATCAGCGGTTCTCTAGG 300

RESULT 11
LOCUS BG707608 551 bp mRNA linear EST 07-MAY-2001
DEFINITION 602670672P1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793704 5', mRNA sequence.
ACCESSION BG707608
VERSION BG707608.1 GI:13984125
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 551)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki Toshioyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLAM10673 row: 0 column: 17
High quality sequence stop: 551.
Location/Qualifiers
1. 551
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4793704"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/note="Organ: brain; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

635 GGGGCGGCAAGCGGAGTCTATGTACACAGCCTTCAAGCGCGCTCGGAGATTCCGGTGGACT 694
6 GGGGCGGCAAGCGGAGTCTATGTACACAGCCTTCAAGCGCGCTCGGAGATTCCGGTGGACT 65
695 ACGTGCGAGTCTGTGACTCGGACCAAGTTGGACCCCATGGCATCTGCTGAGCTCGTGC 754
66 ACGTGCGAGTCTGTGACTCGGACCAAGTTGGACCCCATGGCATCTGCTGAGCTCGTGC 125
755 GGGTACTGCGAGGACCCCGGCTAGGGCTGTGGTGGGAGCTGCGGATCCTTAACC 814
126 GGGTACTGCGAGGACCCCGGCTAGGGCTGTGGTGGGAGCTGCGGATCCTTAACC 185

Query Match 27.2%; Score 291.8; DB 4; Length 551;
Best Local Similarity 99.3%; Pred. No. 1.6e-48;
Matches 293; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 635 GGGGCGGCAAGCGGAGTCTATGTACACAGCCTTCAAGCGCGCTCGGAGATTCCGGTGGACT 694
Db 6 GGGGCGGCAAGCGGAGTCTATGTACACAGCCTTCAAGCGCGCTCGGAGATTCCGGTGGACT 65
Qy 695 ACGTGCGAGTCTGTGACTCGGACCAAGTTGGACCCCATGGCATCTGCTGAGCTCGTGC 754
Db 66 ACGTGCGAGTCTGTGACTCGGACCAAGTTGGACCCCATGGCATCTGCTGAGCTCGTGC 125
Qy 755 GGGTACTGCGAGGACCCCGGCTAGGGCTGTGGTGGGAGCTGCGGATCCTTAACC 814
Db 126 GGGTACTGCGAGGACCCCGGCTAGGGCTGTGGTGGGAGCTGCGGATCCTTAACC 185

ORIGIN

Query Match 25.0%; Score 268.2; DB 9; Length 508;
Best Local Similarity 91.1%; Pred. No. 9e-44;
Matches 285; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 397 GTGTGGATGGCAACCGCGCGAGGACCTTACATGGTTCGACATGTTCCGCGAGGTCTTC 456
Db 1 GTGTGGACGCGCAACCGCGCGAGGACCTTACATGGTGGACATGTTCCGCGAGGTCTTC 60

457 GCTCAGCAGGACCCCGCCACGCTAGTGTGGACCGCAACTACCACCGCTTGGGAACCC 516
61 GCCGACGAGGACCCCGGCGACCTAGTGTGGACCGCAACTACCACCGCTTGGGAGCG 120
517 GCGCGCGCGCGCGGTGGCGCGCGAGCCCTATCGGGAGGTGGAGCGCGAGGATCCTGG 576
121 GCGCGCGCGCGCGCGCGCGCGCGCTTACCGGAGGTGGAGCGCGAAGACCCCGG 180
577 CGGTGGCAGTGGAGGCGCTGCTGAGGACTCGCAGTTCGCTGCTGGTGGCGAGCGCTGG 636
181 CGGTGGCGGTGGAGGCGCTGCTGAGGACGCAAGTTCGCTGCTGGTGGCGAGCGCTGG 240
637 GCGCGAAGCGCGAGGTTCATGTACACAGCCTTCAAGCGCTTCGAGATTCCTGGACTAC 696
241 GCGCGAAGCGCGAGGTTCATGTACACAGCCTTCAAGCGCTTCGAGATTCCTGGACTAC 300
697 GTGAGGTCTGTG 709

Qy

815 CTCTGGACTCTCGGTACGCTTCCTTAAGCAGCCTCGGATCTGGGTAGCCTTCAATGTGG 874
186 CTCTGGACTCTCGGTACGCTTCCTTAAGCAGCCTCGGATCTGGGTAGCCTTCAATGTGG 245

875 AGCGGGCTTGTACAGACTTCTTCCACTGTGTATCTCGCATCAGCGGTTCTCTAGG 929
246 AGCGGGCTTGTACAGACTTCTTCCACTGTGTATCTCGCATCAGCGGTTCTCTAGG 300

RESULT 12
LOCUS CE377866 508 bp DNA linear GSS 27-SEP-2003
DEFINITION tigr-gss-dog-17000362008906 Dog Library Canis familiaris genomic, genomic survey sequence.
ACCESSION CE377866
VERSION CE377866.1 GI:36605926
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 508)
AUTHORS Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K., Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and Venter, J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
MEDLINE 22875432
PubMed 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirkness@tigr.org
Class: shotgun.
Location/Qualifiers
1. 508
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BatXI; Libraries were prepared from peripheral blood"

ORIGIN

Query Match 25.0%; Score 268.2; DB 9; Length 508;
Best Local Similarity 91.1%; Pred. No. 9e-44;
Matches 285; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy 397 GTGTGGATGGCAACCGCGCGAGGACCTTACATGGTTCGACATGTTCCGCGAGGTCTTC 456
Db 1 GTGTGGACGCGCAACCGCGCGAGGACCTTACATGGTGGACATGTTCCGCGAGGTCTTC 60

457 GCTCAGCAGGACCCCGCCACGCTAGTGTGGACCGCAACTACCACCGCTTGGGAACCC 516
61 GCCGACGAGGACCCCGGCGACCTAGTGTGGACCGCAACTACCACCGCTTGGGAGCG 120
517 GCGCGCGCGCGCGGTGGCGCGCGAGCCCTATCGGGAGGTGGAGCGCGAGGATCCTGG 576
121 GCGCGCGCGCGCGCGCGCGCGCTTACCGGAGGTGGAGCGCGAAGACCCCGG 180
577 CGGTGGCAGTGGAGGCGCTGCTGAGGACTCGCAGTTCGCTGCTGGTGGCGAGCGCTGG 636
181 CGGTGGCGGTGGAGGCGCTGCTGAGGACGCAAGTTCGCTGCTGGTGGCGAGCGCTGG 240
637 GCGCGAAGCGCGAGGTTCATGTACACAGCCTTCAAGCGCTTCGAGATTCCTGGACTAC 696
241 GCGCGAAGCGCGAGGTTCATGTACACAGCCTTCAAGCGCTTCGAGATTCCTGGACTAC 300
697 GTGAGGTCTGTG 709

Qy

```
Db 301 GTGCAGGTGAGTG 313
|||||
RESULT 13
BX327795 610 bp mRNA linear EST 08-APR-2004
LOCUS
DEFINITION BX327795 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1027YJ002 5-PRIME, mRNA sequence.
ACCESSION BX327795
VERSION BX327795.2 GI:46277978
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 610)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT On May 2, 2003 this sequence version replaced gi:30340686.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9502.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?8=CS0BAG028ZH06_CS02640_1kc=9502.r

FEATURES
source
Location/Qualifiers
1. 610
/mol_type="mRNA"
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clones="CS0D1027YJ02"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 24.2%; Score 258.8; DB 5; Length 610;
Best Local Similarity 97.0%; Pred. No. 7e-42; Mismatches 0; Indels 0; Gaps 0;
Matches 263; Conservative 0;

Qy 659 ACACAGCGCTTCGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACA 718
|||||
Db 1 ACATGCTTTAAGGCGCTCGGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACA 60
|||||

Qy 719 CAAGTTGGACCCCATGGCACTGTGTGAGTCTGTGCGGGTACTGGAACGAGACCCCCGGG 778
|||||
Db 61 CAAGTTGGACCCCATGGCACTGTGTGAGTCTGTGCGGGTACTGGAACGAGACCCCCGGG 120
|||||

Qy 779 TAGGGCTGTGTGGGAGCTGGGATCTTACCTCTGGACTCTGGGTGAGCTTCC 838
|||||
Db 121 TAGGGCTGTGTGGGAGCTGGGATCTTACCTCTGGACTCTGGGTGAGCTTCC 180
|||||

Qy 839 TAAGCAGCTCGGATCTGGGTAGCTTCAATGTGGAGCGGGCTTGTTCAGAGCTTCTCC 898
|||||
Db 181 TAAGCAGCTCGGATCTGGGTAGCTTCAATGTGGAGCGGGCTTGTTCAGAGCTTCTCC 240
|||||

Qy 899 ACTGTGTATCTGCATCAGCGGTTCTTAGG 929
|||||
Db 241 ACTGTGTATCTGCATCAGCGGTTCTTAGG 271
|||||

RESULT 14
CF239758 894 bp mRNA linear EST 05-AUG-2003
LOCUS
DEFINITION AGENCOURT_15100464 NICHD_XGC_Emb6 Xenopus tropicalis cDNA clone
IMAGE:6994934 5', mRNA sequence.
CF239758
ACCESSION CF239758.1 GI:33442966
VERSION
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 894)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1A014677 row: m column: 13
High quality sequence stop: 699.

FEATURES
source
Location/Qualifiers
1. 894
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:6994934"
/tissue_type="neurala"
/dev_stage="embryo, stages 14-19"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD_XGC Emb6"
/notes="Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 2.1 kb. Constructed by Invitrogen. Note: This is a
Xenopus Gene Collection (XGC) library."

ORIGIN
Query Match 23.9%; Score 256.2; DB 7; Length 894;
Best Local Similarity 63.6%; Pred. No. 2.3e-41;
Matches 409; Conservative 0; Mismatches 228; Indels 6; Gaps 1;

Qy 289 GTGGCGGTGACCATCTCCGCTACAGGAGACCCCGCGTACCTGCGCGAGTGGCTGGCG 348
|||||
Db 54 GTGGCGGTGACCATCTCCGCTACAGGAGACCCCGCGTACCTGCGCGAGTGGCTGGCG 113
|||||

Qy 349 TCCGCGCGCGCTGCTGTACCGCGCGCGGCTGCGCTCATGCTCATGCTGGTGGTGGTGC 408
|||||
Db 114 TCCGCGCGGTGCTGTGAGTACCCCAAGGACAACTCAAGATCATTTGGTCATCGACGG 173
|||||

Qy 409 AACCGCGCGCGGAGACCTCTACATGTCGACATGTTCCGCGAGGTCCTTCGCTGACGAGAC 468
|||||
Db 174 AACCGCGGAGTACGCTCATGATGATGATGATGATGATGATGATGATGATGATGATG 233
|||||

Qy 469 CCGCGCACGTCGTGTGGGACGGCACTACACAGCCCTGGGAGACCCCGCGCGCGCGGGC 528
|||||
Db 234 GTGGGTACCTACGTCGTGGAGGGGCACTACACACTGGCGTAAAGGAGAGACCAAGATGC 293
|||||

Qy 529 GCGGTGGCGCGGAGCTATCGGAGGTGAGCGGAGGATCCCTGGCGCGCTGGCAGTG 588
|||||
Db 294 TCTGTCTGAGGTGTCGAAGCCCTGAATGAAGATGAG-----GGGATACGTATAGTG 347
|||||

Qy 589 GAGGCGCTGTGAGGACTCGCAGGTGCTGTGCTGGTGGCGAGCGCTCGGGGCGGCAAGCGC 648
|||||
```


Db 348 GAGGAACATGGTCAGAACCAAGAGGTGTGTGTCATCATGACGACGTCGGGCGGCAACGG 407
 Qy 649 GAGGTCAATGTACACAGCCCTTCAAGCGCTTCGAGAGATTGCGGTGAGTACGTGACGAGTCTGT 708
 Db 408 GAGGTCAATGTACACAGCCATTCGGGCCATTCGGGACCAACCATGAGTACGTACAGTCTGT 467
 Qy 709 GACTCGGACACAGTTTGGACCCCATGCACTGCTGGAGCTGCTCGGGGTACTGACACGAG 768
 Db 468 GACTCCGACACCAAACTGGACGAACTGGCCACGGTGGAAATGGTGAAGGTCTCTAGAGGCC 527
 Qy 769 GACCCCGGGTAGGGGCTGTGGTGGGACGCTGGGATCCTTAACCCCTCGACTCCCTGG 828
 Db 528 AACGAGCTGTGGGGGCGGTGGAGGAGACGTTCGCATCCTGAACCCCTACGACTCTTC 587
 Qy 829 GTCAGCTCTCAAGCAGCTCGTACTGCTGAGTCTTCAATGTGGAGCGGGCTTGTGAG 888
 Db 588 ATCAGTTTCATGAGCAGCTAGCTTATTGGATGGGTTTAACGTGGAGAGGCGCTGCCAG 647
 Qy 889 AGCTACTTCCACTGTGTATCCTGCATCAGCGGTTCTTAGGTA 931
 Db 648 TCCTACTTCGACTGCGTCTCTGTATCAGTGGCCCCCTGGGAA 690

RESULT 15

BI753116 782 bp mRNA linear EST 25-SEP-2001
 LOCUS 603025961F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196399 5',
 mRNA sequence.

ACCESSION BI753116
 VERSION BI753116.1 GI:15744694

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 782)

TITLE NIH-MGC http://mgs.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LHAM1491 row: j column: 16

High quality sequence stop: 780.

Location/Qualifiers

FEATURES

source

1. 782
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5196399"
 /lab_host="DH10B"
 /clone_lib="NTH_MGC_114"
 /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH_MGC Library."

ORIGIN

Query Match 23.9%; Score 255.8; DB 4; Length 782;
 Best Local Similarity 99.2%; Pred. No. 2.8e-41;
 Matches 257; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 671 AGGCGCTCGGAGATTGCGTGGACTACGTGCAAGGTCTGTGACTCGGACACAAGGTTGGACC 730
 Db 6 AAGCGCTCGGAGATTGCGTGGACTACGTGCAAGGTCTGTGACTCGGACACAAGGTTGGACC 65
 Qy 731 CCATGGCACTGCTGGAGCTCGTGGGTTACTGGACGAGACCCCGGGTAGGGGCTGTG 790
 Db 66 CCATGGCACTGCTGGAGCTCGTGGGTTACTGGACGAGACCCCGGGTAGGGGCTGTG 125
 Qy 791 GTGGGGACGTGCGGATCCTTAACCTCTGGACTCCTGGGTGAGCTTCTTAAGCAGCCTGC 850
 Db 126 GTGGGGACGTGCGGATCCTTAACCTCTGGACTCCTGGGTGAGCTTCTTAAGCAGCCTGC 185
 Qy 851 GATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCACTGTGTATCCT 910
 Db 186 GATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCACTGTGTATCCT 245
 Qy 911 GCATCAGCGGTTCTCTAGG 929
 Db 246 GCATCAGCGGTTCTCTAGG 264

Search completed: March 13, 2005, 11:54:55
 Job time : 4422.94 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 01:47:21 ; Search time 696.034 Seconds
(without alignments)
9108.809 Million cell updates/sec

Title: US-10-672-399-3
Perfect score: 1071
Sequence: 1 atgagacagcaggacgccc.....ccggaccatcgctggatga 1071

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	927.4	86.6	1737	10 ADL13691	ADL13691 Osteoarth
2	927.4	86.6	2088	10 ADL13690	ADL13690 Osteoarth
3	927.4	86.6	2116	2 AAT99541	AAT99541 Human hya
4	925.8	86.4	1737	10 ADL13692	ADL13692 Osteoarth
5	925.8	86.4	2087	10 ADL13694	ADL13694 Osteoarth
6	905.4	84.5	2117	2 AAT96713	AAT96713 Human hya
7	697.2	65.1	231222	10 ADL13693	ADL13693 Osteoarth
8	678.2	63.3	1752	8 ABZ76734	ABZ76734 Mouse hya
9	678.2	63.3	1752	10 AAD59442	AAD59442 Mouse hya
10	678.2	63.3	2102	2 AAT91655	AAT91655 Mouse hya
11	678.2	63.3	2102	2 AAZ10862	AAZ10862 Hyaluron
12	678.2	63.3	2102	3 AAZ288199	AAZ288199 Mouse hya
13	678.2	63.3	2102	3 AAA399987	AAA399987 Murine HA
14	422	39.4	662	6 ABQ27384	ABQ27384 Oligonuc1
15	422	39.4	662	6 ABQ27385	ABQ27385 Oligonuc1
16	393.8	36.8	662	6 ABQ27382	ABQ27382 Oligonuc1
17	393.8	36.8	662	6 ABQ27383	ABQ27383 Oligonuc1
18	338.6	31.6	490	9 ACH39998	ACH39998 Human toe
19	290.6	27.1	1783	10 ADC49212	ADC49212 Rabbit hy
20	272.8	25.5	1662	8 ACC51023	ACC51023 Human bla

21	272.8	25.5	1662	8 ABX76415	ABX76415 Lung canc
22	272.8	25.5	1662	11 ADN39051	ADN39051 Cancer/an
23	272.8	25.5	4049	8 ABX76214	ABX76214 Lung canc
24	272.8	25.5	4049	12 ADN05115	ADN05115 Antipsori
25	272.8	25.5	4220	12 ADJ74905	ADJ74905 Marker ge
26	248.4	23.2	1767	10 ADD93927	ADD93927 Xenopus 1
27	236.2	22.1	1665	2 AAV18822	AAV18822 Murine hy
28	236.2	22.1	1665	3 AAZ88201	AAZ88201 Mouse hya
29	236.2	22.1	1665	8 ABZ76736	ABZ76736 Mouse hya
30	236.2	22.1	1665	10 AAD59444	AAD59444 Mouse hya
31	236.2	22.1	5919	12 ADJ75814	ADJ75814 Marker ge
32	207.4	19.4	1229	12 ADJ74906	ADJ74906 Marker ge
33	176.2	16.5	1653	6 ABL41013	ABL41013 Murine hy
34	176.2	16.5	1659	8 ABZ76735	ABZ76735 Mouse hya
35	176.2	16.5	1659	10 AAD59443	AAD59443 Mouse hya
36	176.2	16.5	2947	2 AAV18821	AAV18821 Murine hy
37	176.2	16.5	2948	6 ABI98697	ABI98697 Mouse isc
38	176.2	16.5	4194	9 ADA49685	ADA49685 DNA encod
39	176.2	16.5	4194	10 ADC59298	ADC59298 Mouse HAS
40	162.6	15.2	2890	3 AAZ88200	AAZ88200 Mouse hya
41	148.8	13.9	1659	10 ADL13696	ADL13696 Osteoarth
42	148.8	13.9	3003	6 ABK84373	ABK84373 Human CDN
43	148.8	13.9	3003	9 ADA49683	ADA49683 DNA encod
44	148.8	13.9	3003	10 ADC59296	ADC59296 Human HAS
45	148.8	13.9	3003	10 ADD12718	ADD12718 Human CDN

ALIGNMENTS

RESULT 1
ADL13691
ID ADL13691 standard; DNA; 1737 BP.
XX
AC ADL13691;
XX
DT 06-MAY-2004 (first entry)
XX
DE Osteoarthritis-associated polymorphic nucleotide #223.
XX
KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN WO2003054166-A2.
XX
PD 03-JUL-2003.
XX
PF 19-DEC-2002; 2002WO-US041225.
XX
PR 20-DEC-2001; 2001US-0342603P.
XX
(INCY-) INCYTE GENOMICS INC.
XX
Jones KA, Schafer A;
XX
WPI; 2003-559141/52.
XX
Determining susceptibility of an individual to joint space narrowing,
osteophyte development and/or joint pain comprises identifying whether
the individual has at least one polymorphism in a polynucleotide encoding
a protein.
XX
Disclosure; SEQ ID NO 223; 297pp; English.
XX
The invention relates to a method of determining susceptibility of an
individual to joint space narrowing and/or osteophyte development and/or
joint pain comprising identifying whether the individual has at least one
polymorphism in a polynucleotide encoding at least one of the protein
listed in the specification. The methods, composition and agent are
useful for modulating the susceptibility of an individual to joint space

CC narrowing and/or osteophyte development and/or joint pain that is
CC associated with a disease, preferably osteoarthritis. The cell line and
CC the non-human animal are useful for screening for an agent for diagnosing
CC an individual having susceptibility to joint space narrowing and/or
CC osteophyte development and/or joint pain. This sequence corresponds to
CC the polynucleotide encoding a protein listed in the specification. (Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences).

XX
SQ Sequence 1737 BP; 232 A; 571 C; 603 G; 331 T; 0 U; 0 Other;

Query Match 86.6%; Score 927.4; DB 10; Length 1737;
Best Local Similarity 99.9%; Pred. No. 1.6e-152;
Matches 928; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAGACAGCAGGACGCGCCCAAGCCACTCTCTGACGCGCCGCGCTGCTCGGCTTGCC 60
Db 1 ATGAGACAGCAGGACGCGCCCAAGCCACTCTCTGACGCGCCGCGCTGCTCGGCTTGCC 60

Qy 61 CGAGGGTGCTGAACCATCGCTTGCCCTGCTCATCTCTGAGCCTCATGACCTGGCCCTAC 120
Db 61 CGAGGGTGCTGAACCATCGCTTGCCCTGCTCATCTCTGAGCCTCATGACCTGGCCCTAC 120

Qy 121 GCCGCCGGGGTGCCGCTGCGCTCGATCGCTACGGCTCTCTGGCTTCGGCTCTACGGG 180
Db 121 GCCGCCGGGGTGCCGCTGCGCTCGATCGCTACGGCTCTCTGGCTTCGGCTCTACGGG 180

Qy 181 GCCTTCCTTTTCAGCGCACCTGGTGCGCAGAGCCTCTTCGCGTACTGGAGCACCGGCGG 240
Db 181 GCCTTCCTTTTCAGCGCACCTGGTGCGCAGAGCCTCTTCGCGTACTGGAGCACCGGCGG 240

Qy 241 GTGCGCGCGCGCGCGCGCGCGCTGTGATGACGACACCGCGCGAGTGTGGCGCTGACC 300
Db 241 GTGCGCGCGCGCGCGCGCGCGCTGTGATGACGACACCGCGCGAGTGTGGCGCTGACC 300

Qy 301 ATCTCCGCTTACAGGAGGACCGCGCTACTGCGCAGTCTGCGCTGCGCGCGCGCGCC 360
Db 301 ATCTCCGCTTACAGGAGGACCGCGCTACTGCGCAGTCTGCGCTGCGCGCGCGCGCC 360

Qy 361 CTGCTGTACCGCGCGCGCGCTGCGCTCTCTCATGTGTGGATGGCAACCGCGCGCGAG 420
Db 361 CTGCTGTACCGCGCGCGCGCTGCGCTCTCTCATGTGTGGATGGCAACCGCGCGCGAG 420

Qy 421 GACCTCTACATGTGTGACATGTTGCGAGGTCTTGTGACGAGACCCCGCCACGTAC 480
Db 421 GACCTCTACATGTGTGACATGTTGCGAGGTCTTGTGACGAGACCCCGCCACGTAC 480

Qy 481 GTGTGGGACGCACTACACACCCCTGGGAAACCGCGCGCGCGCGCGCGCGCGCGCC 540
Db 481 GTGTGGGACGCACTACACACCCCTGGGAAACCGCGCGCGCGCGCGCGCGCGCGCC 540

Qy 541 GGAGCCTATCGGAGGTGGAGCGGAGGATCTTGGCGCGCTGGCAGTGGAGCGCTGGTG 600
Db 541 GGAGCCTATCGGAGGTGGAGCGGAGGATCTTGGCGCGCTGGCAGTGGAGCGCTGGTG 600

Qy 601 AGGACTCGAGGTGCGTGTGCGTGGCGAGCGCTGGGCGCGCAAGCGCGAGGTATGTAC 660
Db 601 AGGACTCGAGGTGCGTGTGCGTGGCGAGCGCTGGGCGCGCAAGCGCGAGGTATGTAC 660

Qy 661 ACAGCCTTCAAGGCGCTCGAGATTCGGTGACTACGTGACGAGTCTGTGACTCGGACCA 720
Db 661 ACAGCCTTCAAGGCGCTCGAGATTCGGTGACTACGTGACGAGTCTGTGACTCGGACCA 720

Qy 721 AGGTGGACCCATGCGACTCTGTGAGCTCTGTGGGGTACTGGACGAGACCCCGGGTA 780
Db 721 AGGTGGACCCATGCGACTCTGTGAGCTCTGTGGGGTACTGGACGAGACCCCGGGTA 780

Qy 781 GGGCTGTGTGGGGACGTGCGATCTTAACCTCTGGAATCTCTGGGTGAGTTCTTA 840
Db 781 GGGCTGTGTGGGGACGTGCGATCTTAACCTCTGGAATCTCTGGGTGAGTTCTTA 840

Qy 841 AGCAGCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCAC 900

Db 841 AGCAGCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCAC 900

Qy 901 TGTGTATCTCTGCAATCAGCGGTTCTCTAGG 929

Db 901 TGTGTATCTCTGCAATCAGCGGTTCTCTAGG 929

RESULT 2

ADL13690
ID ADL13690 standard; DNA; 2088 BP.

XX AC ADL13690;
XX DT 06-MAY-2004 (first entry)
XX DE Osteoarthritis-associated polymorphic nucleotide #222.
XX KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX KW joint space narrowing; osteophyte development; joint pain;
XX KW osteoarthritis; SNP; single nucleotide polymorphism.
XX OS Homo sapiens.
XX PN WO2003054166-A2.
XX PD 03-JUL-2003.
XX PF 19-DEC-2002; 2002WO-US041225.
XX PR 20-DEC-2001; 2001US-0342603P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Jones KA, Schafer A;
XX WPI; 2003-559141/52.
XX DR
XX PT Determining susceptibility of an individual to joint space narrowing,
XX PT osteophyte development and/or joint pain comprises identifying whether
XX PT the individual has at least one polymorphism in a polynucleotide encoding
XX PS a protein.
XX PS Disclosure; SEQ ID NO 222; 297pp; English.
XX CC The invention relates to a method of determining susceptibility of an
XX CC individual to joint space narrowing and/or osteophyte development and/or
XX CC joint pain comprising identifying whether the individual has at least one
XX CC polymorphism in a polynucleotide encoding at least one of the protein
XX CC listed in the specification. The methods, composition and agent are
XX CC useful for modulating the susceptibility of an individual to joint space
XX CC narrowing and/or osteophyte development and/or joint pain that is
XX CC associated with a disease, preferably osteoarthritis. The cell line and
XX CC the non-human animal are useful for screening for an agent for diagnosing
XX CC an individual having susceptibility to joint space narrowing and/or
XX CC osteophyte development and/or joint pain. This sequence corresponds to
XX CC the polynucleotide encoding a protein listed in the specification. (Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences).

XX SQ Sequence 2088 BP; 302 A; 652 C; 711 G; 423 T; 0 U; 0 Other;

Query Match 86.6%; Score 927.4; DB 10; Length 2088;

Best Local Similarity 99.9%; Pred. No. 1.6e-152;

Matches 928; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAGACAGCAGGACGCGCCCAAGCCACTCTCTGACGCGCCGCGCTGCTCGGCTTGCC 60
Db 36 ATGAGACAGCAGGACGCGCCCAAGCCACTCTCTGACGCGCCGCGCTGCTCGGCTTGCC 95
Qy 61 CGAGGGTGCTGACCATCGCTTTCGCGCTGCTCATCTGCGGCTCTCATGACCTGGCCCTAC 120

```
Db 96 CGAGGGTGTGACCATCGCCTTCGCCCTGCTCATCTCGGGGCTCATGACCTGGGCGCTTAC 155
Qy 121 GC CGCGGGGTGCGCTCGCCTCCGATCGTACGCGCTCTGGCCCTTCGGCCCTTACGGG 180
Db 156 GC CGCGGGGTGCGCTCGCCTCCGATCGTACGCGCTCTGGCCCTTCGGCCCTTACGGG 215
Qy 181 GCCTTCCTTTGAGCGACCTGGTGGCGAGAGCCTCTTCGCGTACTGTGAGCACCGGCGG 240
Db 216 GCCTTCCTTTGAGCGACCTGGTGGCGAGAGCCTCTTCGCGTACTGTGAGCACCGGCGG 275
Qy 241 GTGGCGGCGGCGCGGCGCGCTGGATGACGACACCGCGCGAGTGTGGCGTGTACC 300
Db 276 GTGGCGGCGGCGCGGCGCGCTGGATGACGACACCGCGCGAGTGTGGCGTGTACC 335
Qy 301 ATCTCCGCTTACAGAGAGACCCCGCTACCTGCGCCAGTGTGCTGGCGTGGCGCGCGCC 360
Db 336 ATCTCCGCTTACAGAGAGACCCCGCTACCTGCGCCAGTGTGCTGGCGTGGCGCGCGCC 395
Qy 361 CTGCTGTACCGCGCGCGCGCTGGCTGCTCATGTGTGGATGGCAACCGCGCGGAG 420
Db 396 CTGCTGTACCGCGCGCGCGCTGGCTGCTCATGTGTGGATGGCAACCGCGCGGAG 455
Qy 421 GACCTCTACATGGTGCATGTTCCGCGAGGTCTTCGCTGACGAGACCCCGCCAGCTAC 480
Db 456 GACCTCTACATGGTGCATGTTCCGCGAGGTCTTCGCTGACGAGACCCCGCCAGCTAC 515
Qy 481 GTGTGGGACGGCACTACACAGCCCTGGGAACCCCGCGCGCGCGCTGGGCGCC 540
Db 516 GTGTGGGACGGCACTACACAGCCCTGGGAACCCCGCGCGCGCGCTGGGCGCC 575
Qy 541 GGAGCCTATCGGAGGTGGAGCGAGATCTTCGGCGGCTGGCAGTGGAGCGCTGTGTG 600
Db 576 GGAGCCTATCGGAGGTGGAGCGAGATCTTCGGCGGCTGGCAGTGGAGCGCTGTGTG 635
Qy 601 AGGACTTCAGAGCCTCGGAGATTCGGTGGACTACGTGAGCTGTGACTCGGACACA 660
Db 636 AGGACTTCAGAGCCTCGGAGATTCGGTGGACTACGTGAGCTGTGACTCGGACACA 695
Qy 661 ACAGCCTTCAGAGCCTCGGAGATTCGGTGGACTACGTGAGCTGTGACTCGGACACA 720
Db 696 ACAGCCTTCAGAGCCTCGGAGATTCGGTGGACTACGTGAGCTGTGACTCGGACACA 755
Qy 721 AGTTTGGACCCCATGCTGCTGAGCTGTGCGGCTGTGAGACGAGACCCCGGGTA 780
Db 756 AGTTTGGACCCCATGCTGCTGAGCTGTGCGGCTGTGAGACGAGACCCCGGGTA 815
Qy 781 GGGCGTGTGTGGGAGCTGGGATCTTAACCTCTGGAATCTGCTGCTGCTGCTGCTTCTTA 840
Db 816 GGGCGTGTGTGGGAGCTGGGATCTTAACCTCTGGAATCTGCTGCTGCTGCTGCTTCTTA 875
Qy 841 AGCAGCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 900
Db 876 AGCAGCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 935
Qy 901 TGTGTATCTGCATCAGCGGTTCCTAGG 929
Db 936 TGTGTATCTGCATCAGCGGTTCCTAGG 964
```

RESULT 3

AAT99541

ID AAT99541 standard; cDNA; 2116 BP.

AC AAT99541;

DT 21-MAY-1998 (first entry)

DE Human hyaluronan synthase cDNA clone 30C.

KW Hyaluronan synthase; HAS gene; human; hyaluronan acid; cell adhesion;
wound healing; vulnary; tissue repair; scar; keloid; therapy; ds.

OS Homo sapiens.

```
XX Key Location/Qualifiers
FH CDS 36..1772
FT /*tag= a
FT polyA_signal 2066..2071
FT /*tag= b
XX WO9740174-A1.
PN 30-OCT-1997.
PD 17-APR-1997; 97WO-US006350.
PF 22-APR-1996; 96US-00635552.
PR (LEUK-) LEUKOSITE INC.
PA Briskin MJ;
PI WPI; 1997-549359/50.
XX P-PSDB; AAW26765.
DR Human hyaluronan synthase - useful for recombinant production of
PT hyaluronic acid for wound healing, tissue repair and reducing
PT hypertrophic scar and keloid formation.
XX Claim 3; Page 36-38; 58pp; English.
PS cDNA clone 30C includes a coding region for hyaluronan synthase (HAS)
XX (see AAW26765), an enzyme involved in the synthesis of hyaluronan
CC (hyaluronic acid) and which has the ability to confer cell adhesion by
CC the lymphocyte receptor CD44. Clone 30C was isolated using an expression
CC cloning system developed to isolate cDNA clones that encode proteins that
CC confer adhesion of the murine T cell lymphoma TK1. A human mesenteric
CC lymph node expression library was constructed that, upon transfection
CC into CHO/P cells, yielded clone 30C that mediated rosetting of TK1 cells
CC to some of the transfectants. The isolated clone can be utilised in a
CC claimed method for producing HAS in host cells. Such host cells are used
CC in a claimed method for the production of hyaluronan. Hyaluronan is
CC useful for wound healing and tissue repair, and can reduce or prevent
CC hypertrophic scars and keloid formation. It is also used in eye surgery
CC as a replacement for vitreous fluid
SQ Sequence 2116 BP; 330 A; 651 C; 712 G; 423 T; 0 U; 0 Other;
```

```
Query Match 85.6%; Score 927.4; DB 2; Length 2116;
Best Local Similarity 99.9%; Pred. No. 1.6e-152;
Matches 928; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGACGCCCGCGCTGCTCGGCGCTGGCC 60
Db 36 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGACGCCCGCGCTGCTCGGCGCTGGCC 95
Qy 61 CGGAGGTGTGTGACCATCGCCTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
Db 96 CGGAGGTGTGTGACCATCGCCTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 155
Qy 121 GCCCGCGGGTGGCGCTGGCCTCCGATCGCTACGCCCTTCGGGCTTCGGGCTCTACCGG 180
Db 156 GCCCGCGGGTGGCGCTGGCCTCCGATCGCTACGCCCTTCGGGCTTCGGGCTCTACCGG 215
Qy 181 GCCTTCCTTTTCAGCGCACCTGTGTGGCGCAGAGCCTCTTCGGGCTACCTGGAGCACCGGCGG 240
Db 216 GCCTTCCTTTTCAGCGCACCTGTGTGGCGCAGAGCCTCTTCGGGCTACCTGGAGCACCGGCGG 275
Qy 241 GTGCGCGCGCGCGCGCGGCGCGCTGTGATGACACCGCGCGAGTGTGGCGCTGAC 300
Db 276 GTGCGCGCGCGCGCGCGGCGCGCTGTGATGACACCGCGCGAGTGTGGCGCTGAC 335
Qy 301 ATCTCCGCTTACAGAGAGACCCCGCTACCTGCGCCAGTGTGCTGGCGTGGCGCGCC 360
Db 336 ATCTCCGCTTACAGAGAGACCCCGCTACCTGCGCCAGTGTGCTGGCGTGGCGCGCC 395
```

Qy 361 CTGCTGTACCCGCGCGCGCTGCGCTCCTCATGTGTGGATGGCAACCGCGCCGAG 420
Dd |||||
Qy 396 CTGCTGTACCCGCGCGCGCTGCGCTCCTCATGTGTGGATGGCAACCGCGCCGAG 455
Dd |||||
Qy 421 GACCTCTACATGGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGTAC 480
Dd |||||
Qy 456 GACCTCTACATGGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGTAC 515
Qy 481 GTGTGGGACGCACTACACAGCCCTGGGAACCCCGCGCGCGCGCGCGGTGGCGCC 540
Dd |||||
Qy 516 GTGTGGGACGCACTACACAGCCCTGGGAACCCCGCGCGCGCGCGGTGGCGCC 575
Qy 541 GGAGCCTATCGGAGGTGGAGCGGAGATCCTGGCGGCTGGCAGTGGAGGCGCTGGTG 600
Dd |||||
Qy 576 GGAGCCTATCGGAGGTGGAGCGGAGATCCTGGCGGCTGGCAGTGGAGGCGCTGGTG 635
Qy 601 AGGACTCGCAGGTGCTGTGGCGCGCAGCGCTGGGCGCGCAAGCGCGAGGTGATGTAC 660
Dd |||||
Qy 636 AGGACTCGCAGGTGCTGTGGCGCGCAGCGCTGGGCGCGCAAGCGCGAGGTGATGTAC 695
Qy 661 ACAGCCTTCAAGGCGCTCGGAGATTCGGTGACTACGTGAGGTCTGTGACTCGGACACA 720
Dd |||||
Qy 696 ACAGCCTTCAAGGCGCTCGGAGATTCGGTGACTACGTGAGGTCTGTGACTCGGACACA 755
Qy 721 AGGTTGGACCCCATGGCACTGCTGAGCTGCTGGGCTGCTGGAGGACCCCGGGTA 780
Dd |||||
Qy 756 AGGTTGGACCCCATGGCACTGCTGAGCTGCTGGGCTGCTGGAGGACCCCGGGTA 815
Qy 781 GGGGCTGTTGGTGGGAGCGTGGGATCCTTAACCTCTGAGACTCCTGGGTGAGTTCCTA 840
Dd |||||
Qy 816 GGGGCTGTTGGTGGGAGCGTGGGATCCTTAACCTCTGAGACTCCTGGGTGAGTTCCTA 875
Qy 841 AGCAGCTGCGATATCGGAGTACCTTCAATGTGAGCGGGCTTGTGAGACTCTTCAC 900
Dd |||||
Qy 876 AGCAGCTGCGATATCGGAGTACCTTCAATGTGAGCGGGCTTGTGAGACTCTTCAC 935
Qy 901 TGTGTATCCTCATCAGCGGTTCTCTAGG 929
Dd |||||
Qy 936 TGTGTATCCTCATCAGCGGTTCTCTAGG 964

RESULT 4

ID ADL13692 standard; DNA; 1737 BP.
XX ADL13692
AC ADL13692;
DT 06-MAY-2004 (first entry)
XX Osteoarthritis-associated polymorphic nucleotide #224.
XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX joint space narrowing; osteophyte development; joint pain;
XX osteoarthritis; SNP; single nucleotide polymorphism.
XX Homo sapiens.
XX WO2003054166-A2.
XX 03-JUL-2003.
XX 19-DEC-2002; 2002WO-US041225.
XX 20-DEC-2001; 2001US-0342603P.
XX (INCY-) INCYTE GENOMICS INC.
XX Jones KA, Schafer A;
XX WPI, 2003-559141/52.
XX Determining susceptibility of an individual to joint space narrowing,
XX osteophyte development and/or joint pain comprises identifying whether

PT the individual has at least one polymorphism in a polynucleotide encoding
PT a protein.
XX Disclosure; SEQ ID NO 224; 297pp; English.
XX The invention relates to a method of determining susceptibility of an
XX individual to joint space narrowing and/or osteophyte development and/or
XX joint pain comprising identifying whether the individual has at least one
XX polymorphism in a polynucleotide encoding at least one of the protein
XX listed in the specification. The methods, composition and agent are
XX useful for modulating the susceptibility of an individual to joint space
XX narrowing and/or osteophyte development and/or joint pain that is
XX associated with a disease, preferably osteoarthritis. The cell line and
XX the non-human animal are useful for screening for an agent for diagnosing
XX an individual having susceptibility to joint space narrowing and/or
XX osteophyte development and/or joint pain. This sequence corresponds to
XX the polynucleotide encoding a protein listed in the specification. (Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences).

SQ Sequence 1737 BP; 232 A; 569 C; 604 G; 332 T; 0 U; 0 Other;

Query Match 86.4%; Score 925.8; DB 10; Length 1737;
Best Local Similarity 99.8%; Pred. No. 3.1e-152;
Matches 927; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGACGCGCGCGCTGCTCGGCTGGCC 60
Dd |||||
Qy 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGACGCGCGCGCTGCTCGGCTGGCC 60
Qy 61 CGGAGGGTGTGACCATCGGCTTGGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
Dd |||||
Qy 61 CGGAGGGTGTGACCATCGGCTTGGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
Qy 121 GCGCGCGGGTGGCGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 180
Dd |||||
Qy 121 GCGCGCGGGTGGCGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGG 180
Qy 181 GCCTTCTTTTACGGCACTGTGGGCGAGAGCTTCTTGGGCTACCTGGAGCACGGGCGG 240
Dd |||||
Qy 181 GCCTTCTTTTACGGCACTGTGGGCGAGAGCTTCTTGGGCTACCTGGAGCACGGGCGG 240
Qy 241 GTGCGCGCGCGCGCGCGCGCGCGCTGCTGATGACCCACCGCGCGAGTGTGGGCTGACC 300
Dd |||||
Qy 241 GTGCGCGCGCGCGCGCGCGCGCGCTGCTGATGACCCACCGCGCGAGTGTGGGCTGACC 300
Qy 301 ATCTCCGCTTACAGGAGGACCCCGCGTACCTGCGCCAGTGTGGCTTGGCTTGGCTTGG 360
Dd |||||
Qy 301 ATCTCCGCTTACAGGAGGACCCCGCGTACCTGCGCCAGTGTGGCTTGGCTTGGCTTGG 360
Qy 361 CTGCTGTACCGCGCGCGCGCGCTGCGGCTGCTCATGTGTGGATGGCAACCGCGCCGAG 420
Dd |||||
Qy 361 CTGCTGTACCGCGCGCGCGCGCTGCGGCTGCTCATGTGTGGATGGCAACCGCGCCGAG 420
Qy 421 GACCTCTACATGTGCGACATGTTCCGCGAGGTCTTCCTGACGAGGACCCCGCCACGTAC 480
Dd |||||
Qy 421 GACCTCTACATGTGCGACATGTTCCGCGAGGTCTTCCTGACGAGGACCCCGCCACGTAC 480
Qy 481 GTGTGGGACGCAACTACCAACCGCTTGGGAACCCCGCGCGCGCGCGGTGGGCGCC 540
Dd |||||
Qy 481 GTGTGGGACGCAACTACCAACCGCTTGGGAACCCCGCGCGCGCGCGGTGGGCGCC 540
Qy 541 GGAGCCTATCGGAGGTGGAGCGGAGGATCTCTGGGCGGCTGGGAGTGGAGGCGCTGGTG 600
Dd |||||
Qy 541 GGAGCCTATCGGAGGTGGAGCGGAGGATCTCTGGGCGGCTGGGAGTGGAGGCGCTGGTG 600
Qy 601 AGGACTCGCAGGTGCTGTGGCGCGCAGCGCTGGGCGCGCAAGCGCGAGGTGATGTAC 660
Dd |||||
Qy 601 AGGACTCGCAGGTGCTGTGGCGCGCAGCGCTGGGCGCGCAAGCGCGAGGTGATGTAC 660
Qy 661 ACAGCCTTCAAGGCGCTCGGAGATTCGGTGGACTACGTGAGGTCTGTGACTCGGACACA 720
Dd |||||

Db 661 ACAGCCTTCAGGCGCTCGGAGATTCTGGTGGACTACGTGCGAGGTCTGTGACTCGGACACA 720
Qy 721 AGTTGGACCCCATGGCACTGCTGAGAGTCTGTCGGGTACTGGACGAGACCCCGGGTA 780
Db 721 AGTTGGACCCCATGGCACTGCTGAGAGTCTGTCGGGTACTGGACGAGACCCCGGGTA 780
Qy 781 GGGGCTGTTGGTGGGAGCGTGGGATCCTTTAAACCTCTGGACTCTGGGTGAGCTTCTTA 840
Db 781 GGGGCTGTTGGTGGGAGCGTGGGATCCTTTAAACCTCTGGACTCTGGGTGAGCTTCTTA 840
Qy 841 AGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 900
Db 841 AGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 900
Qy 901 TGTGTATCTGCATCAGCGGTTCTCTAGG 929
Db 901 TGTGTATCTGCATCAGCGGTTCTCTAGG 929

RESULT 5

ID ADL13694 standard; DNA; 2087 BP.
XX
AC ADL13694;
XX
DT
XX
DE
XX
KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
FN WO2003054166-A2.
XX
PD
XX
PP 19-DEC-2002; 2002WO-US041225.
XX
PR 20-DEC-2001; 2001US-0342603P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Jones KA, Schafer A;
XX
XX WPI; 2003-559141/52.

Determining susceptibility of an individual to joint space narrowing, osteophyte development and/or joint pain comprises identifying whether the individual has at least one polymorphism in a polynucleotide encoding a protein.

Disclosure; SEQ ID NO 226; 297pp; English.

The invention relates to a method of determining susceptibility of an individual to joint space narrowing and/or osteophyte development and/or joint pain comprising identifying whether the individual has at least one polymorphism in a polynucleotide encoding at least one of the protein listed in the specification. The methods, composition and agent are useful for modulating the susceptibility of an individual to joint space narrowing and/or osteophyte development and/or joint pain that is associated with a disease, preferably osteoarthritis. The cell line and the non-human animal are useful for screening for an agent for diagnosing an individual having susceptibility to joint space narrowing and/or osteophyte development and/or joint pain. This sequence corresponds to the polynucleotide encoding a protein listed in the specification. (Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences).

Sequence 2087 BP; 302 A; 650 C; 712 G; 423 T; 0 U; 0 Other;

Query Match 86.4%; Score 925.8; DB 10; Length 2087;
Best Local Similarity 99.8%; Pred. No. 3.1e-152;
Matches 927; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ATGAGACGACGAGGAGCGCCCAAGCCCACTCTGACGCGCGCGCTGCGGCTGCGCC 60
Db 36 ATGAGACGACGAGGAGCGCGCCCAAGCCCACTCTGACGCTGCGGCTGCGGCTGCGCC 95
Qy 61 CGGAGGGTGTGACCATCGCCTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
Db 96 CGGAGGGTGTGACCATCGCCTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 155
Qy 121 GCCGCGCGGGTGGCGCTTCCGATCGCTACCGCTTCTGGGCTTCTGGGCTTCTACCGG 180
Db 156 GCCGCGCGGGTGGCGCTTCCGATCGCTACCGCTTCTGGGCTTCTGGGCTTCTACCGG 215
Qy 181 GCCTTCTTTCAGCGCACCTGGTGGCGCAGAGCTTCTGGGCTTCTGGGCTTCTGGGCGG 240
Db 216 GCCTTCTTTCAGCGCACCTGGTGGCGCAGAGCTTCTGGGCTTCTGGGCTTCTGGGCGG 275
Qy 241 GTGGCGCGCGCGCGCGCGCGCTGATGACGACACCGCGCGCAGTGTGGCGCTGACC 300
Db 276 GTGGCGCGCGCGCGCGCGCGCTGATGACGACACCGCGCGCAGTGTGGCGCTGACC 335
Qy 301 ATCTCCGCTTACCAGGAGGACCCCGCGCTTACCTGGCGCAGTGCCTGGCGTCCCGCGGCC 360
Db 336 ATCTCCGCTTACCAGGAGGACCCCGCGCTTACCTGGCGCAGTGCCTGGCGTCCCGCGGCC 395
Qy 361 CTGCTGTATACCGCGCGCGCGCTGCGCTTCTCATGTGTGGATGGCAACCGCGCGCGAG 420
Db 396 CTGCTGTATACCGCGCGCGCGCTGCGCTTCTCATGTGTGGATGGCAACCGCGCGCGAG 455
Qy 421 GACCTCTACATGTGTCGACATGTCGCGGAGTCTTCGCTGACGAGGACCCCGCACGTCAC 480
Db 456 GACCTCTACATGTGTCGACATGTCGCGGAGTCTTCGCTGACGAGGACCCCGCACGTCAC 515
Qy 481 GTGTGGGACGCGCAACTACACAGCCCTGGGAAACCCCGCGCGCGCGCGCGCTGGGCGCC 540
Db 516 GTGTGGGACGCGCAACTACACAGCCCTGGGAAACCCCGCGCGCGCGCGCTGGGCGCC 575
Qy 541 GGAGCCTATCGGAGGTGGAGCGGAGGATCCTGGGCGGCTGGCACTGGAGGCGCTGGTG 600
Db 576 GGAGCCTATCGGAGGTGGAGCGGAGGATCCTGGGCGGCTGGCACTGGAGGCGCTGGTG 635
Qy 601 AGGACTCGCAGTGTGCTGCGCGCAGCGCTGGGCGCGCAAGCGGAGGTCATGTAC 660
Db 636 AGGACTCGCAGTGTGCTGCGCGCAGCGCTGGGCGCGCAAGCGGAGGTCATGTAC 695
Qy 661 ACAGCCTTCAAGGCGCTCGGAGATTTCGCTGAGCTACGTGACAGGTCTGTGACTCGGACACA 720
Db 696 ACAGCCTTCAAGGCGCTCGGAGATTTCGCTGAGCTACGTGACAGGTCTGTGACTCGGACACA 755
Qy 721 AGGTGGACCCCATGGCACTGCTGAGAGTCTGTCGGGTACTGGACGAGACCCCGGGTA 780
Db 756 AGGTGGACCCCATGGCACTGCTGAGAGTCTGTCGGGTACTGGACGAGACCCCGGGTA 815
Qy 781 GGGGCTGTTGGTGGGAGCGTGGGATCCTTAACCTCTGGACTCTGGGCTGAGCTTCTTA 840
Db 816 GGGGCTGTTGGTGGGAGCGTGGGATCCTTAACCTCTGGACTCTGGGCTGAGCTTCTTA 875
Qy 841 AGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 900
Db 876 AGCAGCCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCAC 935
Qy 901 TGTGTATCTGCATCAGCGGTTCTCTAGG 929
Db 936 TGTGTATCTGCATCAGCGGTTCTCTAGG 964

RESULT 6

AAT96713
ID AAT96713 standard; DNA; 2117 BP.

XX AAT96713;
XX AC
XX 22-APR-1998 (first entry)
XX DT
XX DE
XX KW Human hyaluronate synthetase coding sequence.
XX KW Hyaluronate synthetase; human; hyaluronic acid; drug preparation;
XX KW cosmetic preparation; gene therapy; carcinogenesis; ss.
XX OS
XX XX Homo sapiens.
XX FH Key Location/Qualifiers
XX FT CDS 149..1780
XX FT /*tag= a
XX XX
XX PN WO9738113-A1.
XX PD 16-OCT-1997.
XX XX 31-MAR-1997; 97WO-JP001111.
XX XX 05-APR-1996; 96JP-00084326.
XX PR 30-APR-1996; 96JP-00109663.
XX XX (SEBK) SEIKAGAKU CORP.
XX PA
XX XX Itano N, Kimata K;
XX PI
XX DR WPI; 1997-512726/47.
XX DR P-PSDB; AAW36503.
XX XX
XX PT DNA encoding human hyaluronate synthetase - for industrial scale
XX PT production of hyaluronic acid used in generating anti-carcinogenic drugs
XX PT or for cosmetics.
XX XX
XX PS Claim 3; Page 23-27; 35pp; Japanese.
XX XX
XX CC This sequence encodes a human hyaluronate synthetase, and is the coding
XX CC sequence of the invention. The encoded enzyme is useful for industrial
XX CC scale production of hyaluronic acid for use in the preparation of drugs
XX CC and cosmetics. The drugs can also be used in compositions for the
XX CC treatment of disorders involving the lowering of hyaluronic acid
XX CC production. The peptides may be used for the preparation of antibodies
XX CC recognising hyaluronate synthetase, e.g. for diagnostic purposes.
XX CC Antisense DNA or RNA corresponding to the DNA can be used in gene therapy
XX CC treatment of carcinogenesis
XX XX
XX SQ Sequence 2117 BP; 306 A; 673 C; 703 G; 435 T; 0 U; 0 Other;

Query Match 84.5%; Score 905.4; DB 2; Length 2117;
Best Local Similarity 98.8%; Pred. No. 1.1e-148;
Matches 912; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 7 CAGCAGGACGGCCCAAGCCACTCTCGACGCCCGCTGCTCGGCTCGCCGAGG 66
Db 50 CAGCAGGACGGCCCAAGCCACTCTCGACGCCCGCTGCTCGGCTCGCCGAGG 109
Qy 67 GTGCTGACCAATCGCTTCGCCCTGCTCATCTCGGCTCATGACCTGGGCTACGCCGC 126
Db 110 GTGCTGACCAATCGCTTCGCCCTGCTCATCTCGGCTCATGACCTGGGCTACGCCGC 169
Qy 127 GGGGTGCGGCTGGGCTTCGGATCGCTACGGCTCTCTGGCTTCGGGCTTCAGGGGCTTC 186
Db 170 GGGGTGCGGCTGGGCTTCGGATCGCTACGGCTCTCTGGCTTCGGGCTTCAGGGGCTTC 229
Qy 187 CTTTCAGCGCACCTGGTGGCGCAGAGCTCTTCGGTACCTGGAGCACCGCGGGTGGCG 246
Db 230 CTTTCAGCGCACCTGGTGGCGCAGAGCTCTTCGGTACCTGGAGCACCGCGGGTGGCG 289
Qy 247 GCGCGGCGCGGGGCGCGCTGGATGACGCCACCGCGCGCAGTGTGGGCTGACCATCTCC 306
Db 290 GCGCGGCGCGGGGCGCGCTGGATGACGCCACCGCGCGCAGTGTGGGCTGACCATCTCC 349

Qy 307 GCCTACCAAGGAGGACCCCGGCTACCTGGCGCAGTGCCTGGCGTCCGCCCGCCCTGCTG 366
Db 350 GCCTACCAAGGAGGACCCCGGCTACCTGGCGCAGTGCCTGGCGTCCGCCCGCCCTGCTG 409
Qy 367 TACCCGCGCGCGCGCTGCGGCTCCTCATGTGTGGATGGCAACCCGCCGCGAGACCTC 426
Db 410 TACCCGCGCGCGCGCTGCGGCTCCTCATGTGTGGATGGCAACCCGCCGCGAGACCTC 469
Qy 427 TACATGTCGACATGTTCCCGGAGGTTCTTGCTGACGAGGACCCCGCGCATGCTGTGG 486
Db 470 TACATGTCGACATGTTCCCGGAGGTTCTTGCTGACGAGGACCCCGCGCATGCTGTGG 529
Qy 487 GACGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGCGGCGCGCGCGAGCC 546
Db 530 GACGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGCGGCGCGCGAGCC 589
Qy 547 TATCGGAGGTGGAGGCGGAGGATCCTGGGCGGCTGGCAGTGGAGGCGCTGCTGAGGACT 606
Db 590 TATCGGAGGTGGAGGCGGAGGATCCTGGGCGGCTGGCAGTGGAGGCGCTGCTGAGGACT 649
Qy 607 CGCAGGTGCGTGTGCTGGCGCAGCGCTGGGCGCGGCAAGCGCGAGGTCATGTACAGCC 666
Db 650 CGCAGGTGCGTGTGCTGGCGCAGCGCTGGGCGCGGCAAGCGCGAGGTCATGTACAGCC 709
Qy 667 TTCAAGGCGCTCGGAGATTCCGTTGGACTACGTGCGAGGTCTGTGACTCGGACACAAAGTTG 726
Db 710 TTCAAGGCGCTCGGAGATTCCGTTGGACTACGTGCGAGGTCTGTGACTCGGACACAAAGTTG 769
Qy 727 GACCCCATGCGACTGCTGGAGCTCGTGGGCTACTGGACGAGGACCCCGGCTAGGGGCT 786
Db 770 GACCCCATGCGACTGCTGGAGCTCGTGGGCTACTGGACGAGGACCCCGGCTAGGGGCT 829
Qy 787 GTTGGTGGGACGTCGCGATCCTTAACCTCTGAGCTCTCTGGGTCAGCTTCTTAAGCAGC 846
Db 830 GTTGGTGGGATGTCGCGATCCTTAACCTCTGAGCTCTCTGGGTCAGCTTCTTAAGCAGC 889
Qy 847 CTGCGATACCTGGGTAGGCTTCAATGTGGAGGCGGCTTGTGAGAGCTACTTCCACTGTGTA 906
Db 890 CTGCGATACCTGGGTAGGCTTCAATGTGGAGGCGGCTTGTGAGAGCTACTTCCACTGTGTA 949
Qy 907 TCCTGCAATCAGCGGTTCTCTAGG 929
Db 950 TCCTGCAATCAGCGGTTCTCTAGG 972

RESULT 7
ADL13693/c
ID ADL13693 standard; DNA; 231222 BP.
XX AC ADL13693;
XX DT 06-MAY-2004 (first entry)
XX DE Osteoarthritis-associated polymorphic nucleotide #225.
XX KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX KW joint space narrowing; osteophyte development; joint pain;
XX KW osteoarthritis; SNP; single nucleotide polymorphism.
XX OS Homo sapiens.
XX PN WO2003054166-A2.
XX PD 03-JUL-2003.
XX PR 19-DEC-2002; 2002WO-US041225.
XX PR 20-DEC-2001; 2001US-0342603P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Jones KA, Schafer A;

XX WPI; 2003-559141/52.
XX Determining susceptibility of an individual to joint space narrowing,
XX osteophyte development and/or joint pain comprises identifying whether
XX the individual has at least one polymorphism in a polynucleotide encoding
XX a protein.
XX
XX Disclosure; SEQ ID NO 225; 297pp; English.
XX
XX The invention relates to a method of determining susceptibility of an
XX individual to joint space narrowing and/or osteophyte development and/or
XX joint pain comprising identifying whether the individual has at least one
XX polymorphism in a polynucleotide encoding at least one of the protein
XX listed in the specification. The methods, composition and agent are
XX useful for modulating the susceptibility of an individual to joint space
XX narrowing and/or osteophyte development and/or joint pain that is
XX associated with a disease, preferably osteoarthritis. The cell line and
XX the non-human animal are useful for screening for an agent for diagnosing
XX an individual having susceptibility to joint space narrowing and/or
XX osteophyte development and/or joint pain. This sequence corresponds to
XX the polynucleotide encoding a protein listed in the specification. (Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences).
XX
XX Sequence 231222 BP; 64006 A; 52087 C; 53478 G; 61650 T; 0 U; 1 Other;
XX
XX Query Match 65.1%; Score 697.2; DB 10; Length 231222;
XX Best Local Similarity 99.6%; Pred. No. 1.7e-112;
XX Matches 699; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX 7 CAGCAGGACGCGCCCAAGCCCACTCTCGACGCGCGCTGCTCGGCTCGCGCGAGG 66
XX 149122 CAGCAGGACGCGCCCAAGCCCACTCTCGACGCTGCGCTGCTCGGCTGCGCGAGG 149063
XX
XX 67 GTGCTGACCATCGCCTTCCGCTGCTATCTGCGCTCATGACCTGCGGCTACGCGGCC 126
XX 149062 GTGCTGACCATCGCCTTCCGCTGCTATCTGCGCTCATGACCTGCGGCTACGCGGCC 149003
XX
XX 127 GGGGTGCGCTGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 186
XX 149002 GGGGTGCGCTGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 148943
XX
XX 187 CTTTCAGCGCACCTGGTGGCGCAGAGCCTCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 246
XX 148942 CTTTCAGCGCACCTGGTGGCGCAGAGCCTCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 148883
XX
XX 247 GCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 306
XX 148882 GCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 148823
XX
XX 307 GCCTACCGAGGAGGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 366
XX 148822 GCCTACCGAGGAGGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 148763
XX
XX 367 TACCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 426
XX 148762 TACCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 148703
XX
XX 427 TACATGCTGCAATGTTCCGCGAGGTCTTCGCTGACGAGGACCGCGCGCGCGCGCGCTGCTGCTG 486
XX 148702 TACATGCTGCAATGTTCCGCGAGGTCTTCGCTGACGAGGACCGCGCGCGCGCGCGCTGCTGCTG 148643
XX
XX 487 GACGCGCACTACCAACCGCGCTGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546
XX 148642 GACGCGCACTACCAACCGCGCTGCGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 148583
XX
XX 547 TATCGGAGGTGGAGGCGGAGGATCTCGGCGGCTGCGGAGTGGAGGCGCTGCTGCTGCTGCTGCTG 606
XX 148582 TATCGGAGGTGGAGGCGGAGGATCTCGGCGGCTGCGGAGTGGAGGCGCTGCTGCTGCTGCTGCTG 148523
XX
XX 607 CGCAGGTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 666

Db 148522 CGCAGGTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 148463
Qy 667 TTCAAGGCGCTCGGAGATTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708
Db 148462 TTCAAGGCGCTCGGAGATTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 148421
RESULT 8
ABZ76734
ID ABZ76734 standard; cDNA; 1752 BP.
AC ABZ76734;
DT 01-APR-2000 (first entry)
DE Mouse hyaluronan synthase 1 encoding cDNA SEQ ID NO:1.
XX Mouse; hyaluronan synthase; HAS; HAS1; enzyme; ophthalmological;
XX osteopathic; antiarthritic; gene therapy; angiogenesis inhibitor;
XX hyaluronic acid; dihydrazide; bioconjugate; dry eye syndrome; eye;
XX osteoarthritis; gene; ss.
XX Mus sp.
XX
XX Key Location/Qualifiers
XX CDS 1..1752
XX /*tag= a
XX /product= "hyaluronan synthase 1 (HAS1)"
XX WO2003006068-A1.
XX 23-JAN-2003.
XX 10-JUL-2001; 2001WO-US021785.
XX 10-JUL-2001; 2001WO-US021785.
XX (CLEA-) CLEAR SOLUTIONS BIOTECH INC.
XX Dehazya P, Chen W;
XX WPI; 2003-221664/21.
XX P-PSDB; ABP96028.
XX Novel dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate
XX for treating dry eye syndrome, has derivatized hyaluronic acid
XX crosslinked to nucleic acid encoding protein with hyaluronan synthase
XX activity.
XX Claim 19; Page 56; 62pp; English.
XX
XX The present invention describes a dihydrazide derivatized hyaluronic acid
XX (HA)/nucleic acid (NA) bioconjugate (I), with derivatized HA linked to NA
XX encoding protein with hyaluronan synthase (HAS) activity, where NA has
XX sequence with 79-85% sequence identity (SI) to a sequence of 1752, 1659
XX or 1665 nucleotides (see ABZ76734 to ABZ76736) or a sequence encoding
XX proteins with 66-90% SI to a sequence of 583, 552 or 554 amino acids (see
XX ABP96028 to ABP96030). (I) has ophthalmological, osteopathic and
XX antiarthritic activities, and can be used in gene therapy and as an
XX inhibitor of angiogenesis, and as an inducer of expression of (HA) in
XX human corneal epithelial cell. (I) is useful for transfecting a cell of
XX an eye with nucleic acid, where the nucleic acid comprises (ABZ76734 to
XX ABZ76736), by contacting the cell with (I) comprising the nucleic acid.
XX (I) is useful for treating dry eye syndrome in an individual. (I) is
XX useful in gene therapy applications for the treatment of a variety of
XX medical conditions including dry eye syndrome or other medical conditions
XX where an increase in the production of (HA) in the eye would be
XX therapeutic (e.g., osteoarthritis of the articular joints). (I) is also
XX useful for inhibiting angiogenesis for the treatment of macular
XX degeneration or genes related to lipid biosynthesis that helps to restore
XX the lipid component of the tear film, and as reagents for in vitro
XX transformation of any cell, preferably a eukaryotic cell, more preferably

QY 250 GGGCGCG-----GGGGCGCTGGATGACAGCCACCGCGCGCTGGCG 294
Db |||||
247 GCTGCGCGCGCTCTCTTGGCGAAGGGCGCCCTGGATGCGGCCACTGACAGCGTGGCA 306
QY 295 CTGACCATCTCCGCTTACAGAGGACCCGCGTACCTGGCCAGTGGCTGGGTCGCGC 354
Db |||||
307 CTCACCATCTCAGCTTACCAAGAGGATCCCGCTTACCTGGCGCAGTGTGACCTCCGCG 366
QY 355 CGCGCCCTCTGTATCCCGCGCGCGCTGCGCGTCTCATGTGTGTGATGCCAACCCG 414
Db |||||
367 CGCGCCCTCTGTATCCCGCGCACAGAGTTACGGTGTCTATGGTGTGACGCGAACCCG 426
QY 415 GCGGAGGACCTTACATGTGTGACATGTTCGCGGAGGTCTTCGTGTGACGAGACCCGCG 474
Db |||||
427 GCTGAGGATCTGTATGTGTGACATGTTCGAGAGTCTTCGCGATGAGGACCCCGCG 486
QY 475 AGTGTGTGTGGAGCGGCACTACACAGACCCCTGGACCCCGCGCG---CGGGCGG 531
Db |||||
487 ACTTATGTGTGGATGGCACTACCATCAGCCCTGGGAACCGAGCGAGGCTACGGGCGCT 546
QY 532 GTGGGCGCGGAGCTTATCGGAGGTGAGGGCGGAGATCTTGGGCGGCTGGCAGTGGAG 591
Db |||||
547 GTCGCTGAAGTGGCTTACCGGAGGTGAGGGCGGAGACCCCGGGCGTGGCGGTGGAG 606
QY 592 GCGCTGTGTGAGACTCGCAGGTGCTGTGCTGTGCGGCGACGCTGGGGCGGCAAGCGCGAG 651
Db |||||
607 GCGCTGTGTGAGAACACGCGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 666
QY 652 GTCATGTATACAGCTTCAAGCGCTCGAGATTCGCTGTGCTGTGCTGTGCTGTGCTGTG 711
Db |||||
667 GTCATGTATACAGCTTCAAGCGCATGGGCGGACTCGTGTGCTGTGCTGTGCTGTGCTGTG 726
QY 712 TCGGACACAAAGTTGGACCCCATGTCGCTGTGAGCTGCTGCGGTACTGTGACGAGGAC 771
Db |||||
727 TCAGACACAGACTAGACCCCATGCTGTGAGCTGTGCTGAGTGTGTGATGAGAC 786
QY 772 CCGCGGTAGGGCTGTGTGTGGGACGCTGGGATCTTAAACCTCTGTGACTCTTGGGTC 831
Db |||||
787 CCGCGGTAGGGCTGTGTGTGGGAGGATGTGAGGATCTTAAACCTCTGTGACTCTTGGGTC 846
QY 832 AGCTTCTTAAAGCAGCTCGATCTGCTGTGAGCTTCAATGTGTGAGCGGGCTGTGAGAGC 891
Db |||||
847 AGCTTCTTGTGAGCAGCTTCTGTACTGTGGTGTGCTTCAATGTGTGAGACGAGCTTGTGAGAGC 906
QY 892 TACTTCCACTGTGTCTTGTGATCAGCGGTCTCTTAGGT 930
Db |||||
907 TACTTCCACTGTGTCTTGTGATCAGCGGTCTCTTAGGT 945

RESULT 10
ID AAT91655
XX AAT91655 standard; DNA; 2102 BP.
AC AAT91655;
XX
XX
DT 17-OCT-2003 (revised)
DT 19-DEC-1997 (first entry)
XX
DE Mouse hyaluronate synthase genomic DNA.
XX
KW Hyaluronate synthase; hyaluronic acid; polysaccharide; ds.
XX
OS Mus sp; (strain FM3A).
XX
FH Location/Qualifiers
FT 49..1800
FT /*tag= a
FT /product= "Hyaluronate synthase"
XX
PN JP09224674-A.
XX
XX 02-SEP-1997.
XX

PF 26-FEB-1996; 96JP-00038336.
XX
PR 26-FEB-1996; 96JP-00038336.
XX
PA (KAGG) KAGAKU GIJUTSUCHO CHOKAN KANBO.
XX
XX WPI; 1997-484102/45.
DR P-PSDB; AAW30704.
XX
XX Hyaluronate synthase isolated from mouse cells - useful for large-scale
production of hyaluronic acid.
PS Claim 2; Page 10-13; 15pp; Japanese.
CC A mouse-derived cell strain, FM3A P15-A, was cultured in Eagle minimal
medium containing 10 % heat-inactivated bovine serum, twice concentration
of amino acids and vitamins and penicillin and streptomycin at 37 degrees
Celsius. The culture was subjected to immobilised erythrocyte exclusion
to examine the extent of extracellular formation of hyaluronic acid
matrix. Cells which showed high formation were recovered and named FM3A
HA1. Cells synthesising no hyaluronic acid were recovered and named HAS-
The HAS- cell in which polyoma large T antigen was expressed was prepared
(HAS- P cell). Poly(A)+ RNA was isolated from FM3A HA1 and cDNA was
prepared and was used for constructing a library in HAS- P cells. Cells
having hyaluronic acid synthetic activity were selected from the
transformants and plasmid DNA was recovered and amplified in E.coli. The
resulting genomic DNA sequence codes for hyaluronate synthase having a
sequence of 583 amino acids which is used for large-scale production of
hyaluronic acid, a component of drugs and cosmetics. (Updated on 17-OCT-
2003 to standardise OS field)
SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;
Query Match 63.3%; Score 678.2; DB 2; Length 2102;
Best Local Similarity 84.5%; Pred. No. 3.9e-105;
Matches 793; Conservative 0; Mismatches 128; Indels 18; Gaps 2;
QY 10 CAGGACGCGCCCAAGCCCACTCTCTGACGCGCGCGCTGCTCGGCTGCGCGAGGGTG 69
Db |||||
55 CAGNCATGCCANAGCCCTCAGAGCGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 114
QY 70 CTGACCATGCGCTTGGCGCTGCTCATCTGCGGCTCATGACCTGCGGCTACGCGCGCGG 129
Db |||||
115 CTCAGATCATCTTGGCGCTGCTCATCTGCGGCTCATGACCTGCGGCTACGCGCAGGC 174
QY 130 GTGCGCTGGCTCGGATCGCTACGGCTCTCTGCGCTTGGGCTCTACGCGGCTTCTT 189
Db |||||
175 GTTCTCTGGCTTCAGATCGCTATGGAATCTCTGCGCTTGTGCTCTATGGGCGCATTCCTC 234
QY 190 TCAGCGCACCTGTGGCGCAGAGCTCTTCTGCTACCTGGAGACCGCGGCTGGCGCG 249
Db |||||
235 AGCGCACACCTAGTGGGCACAGAGCTCTTCTGCTTACCTTGGAGACCGAGGCTGGCAGCG 294
QY 250 GCGGCGCG-----GGGGCGCTGGATGACAGCCACCGCGCGCAGTGTGCGG 294
Db |||||
295 GCTGCGCGCGCTCTTGGCGAAGGGCGCCCTGGATGCGGCGCACTGACGCGAGCGTGGCA 354
QY 295 CTGACCATCTCGGCTACAGAGGACCCCGGCTACCTGCGCGAGTGGCTGGCGTCCGCG 354
Db |||||
355 CTCACCATCTCAGGCTTACCAAGAGATCCCGCTTACCTGCGCGCACTGTGTTGACCTCCGCG 414
QY 355 CGCGCCCTGTGTATCCCGCGCGCGCTGCGCGCTGCTCATGGTGTGTGATGGCAACCGC 414
Db |||||
415 CGCGCTTGTGTATCCCGCGCACAGAGTTACGCGTGTCTATGGTGTGACGCGCAACCGC 474
QY 415 GCGGAGGACCTCTATAGTGTGACATGTTCGCGAGGCTTCTCGTGTACGAGACCCCGCG 474
Db |||||
475 GCTGAGGATCTGTATGTGTGACATGTTCGAGAGTCTTTCGCGGATGAGAGACCCCGCG 534
QY 475 ACGTACGTGTGGAGCGGCAACTACCAACGAGCCCTGGGAAACCGCGGCGG---CGGGCGG 531
Db |||||
535 ACTTATGTGTGGATGGCAACTACCATCAGCCCTGGGAAACCGAGCGGAGGCTACGCGGCGT 594

DR P-PSDB; AAB09948.
XX DNA used for gene targeting of the hyaluronate synthase gene (HAS1).
PT
XX Example; Page 7-9; 14pp; Japanese.
PS
XX This invention describes a novel DNA which contains an heterologous DNA
CC (I) to be introduced to chromosomal DNA of a host cell, a first and
CC second homologous region DNA (II) connected respectively to the 5' and 3'
CC sides of the introduced DNA, and a negative marker gene (III) expressable
CC in the host cell. (I) contains a positive marker gene expressable in the
CC host cell. (I), (II) and the region encoding the intracellular loop of
CC the hyaluronate synthase (HAS) 1 protein from chromosomal DNA is
CC constituted so as to cause a homologous recombination. The DNA can be
CC used for gene targeting of the HAS1 gene. This sequence encodes the
CC murine HAS1 protein described in the method of the invention
XX
SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;
Query Match 63.3%; Score 678.2; DB 3; Length 2102;
Best Local Similarity 84.5%; Pred. No. 3.9e-105;
Matches 793; Conservative 0; Mismatches 125; Indels 18; Gaps 2;
QY 10 CAGGACGGCCCAAGCCCACTCTCTGACGCCCGCGCTCTCGGCTTGGCCCGGAGGGTG 69
Db 55 CAGGACATGCCAAGCCCTCAGAGCAGCGGTTGCTCTTGGCTTGGCCCGGAGGAGCA 114
QY 70 CTGACCATCGCTTGGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTACGCGCCGGG 129
Db 115 CTCAGATCATCTTTGGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCGCAGGC 174
QY 130 GTGCCGCTGGCTCGATCGCTACGCGCTCTGGGCTTGGGCTTCTAGCGGCTTCTT 189
Db 175 GTTCTCTGGCTTCAGATCGCTATGGACTCTTGGCTTGGGCTTATGGGCAATTCCTC 234
QY 190 TCAGCGCACTGTGGCGCAGAGCCTCTTCGCTACTTGGAGCACCGCGGGTGGCGCG 249
Db 235 AGCGCACACCTAGTGGCACAGAGCCTCTTCGCTTACTTGGAGCACCGAAGGTTGGCAGC 294
QY 250 GCGGGCG-----GGGCGGCTGGATGAGGACCGCGCGCAGTGTGGCG 294
Db 295 GCTGCGCGGCGCTCTTGGCGAAGGGGCGCTTGGATGCGGCGCTGTCACGCGAGGTTGCA 354
QY 295 CTGACCATCTCGCTACAGAGAGACCGCGCTACCTGCGCCAGTGGCTCGCGTCCGCC 354
Db 355 CTCACCATCTCAGCTTACAGAGATCCCGCTTACTTGGCCAGTGTTCACCTCCGCG 414
QY 355 CGCGCTCTGTATCCCGCGCGCGCTGCGCGCTCTCATGTGGTGGATGGCAACCGC 414
Db 415 CGCGCTTGTGTATCCCGCGCACAGAGTTTACGCGTGTCTCATGTGGTGGAGCGCAACGC 474
QY 415 GCGAGGACCTCTACATGGTGCACATGTTTCGCGAGAGTCTTCGCTGACGAGACCCCGCC 474
Db 475 GCTGAGGATCTGTATAGTGGACATGTTTCGAGAAAGTCTTCGCGGATGAGGACCCCGCC 534
QY 475 ACCTACGTGTGGAGCGCACTTACACAGCCCTGGGAACCGCGGCG---CGGGCGCG 531
Db 535 ACTATGTGTGGAGTGGCACTTACATCAGCCCTTGGGAACCGAGCGAGGCTTACGGGCGCT 594
QY 532 GTGGCGCGCGGAGCTATTCGGAGGTGGAGCGGAGGATCTTCGCGCGCTTGGCAGTGGAG 591
Db 595 GTGCGTGAAGTGCTTACCGGAGGTGGAGCGGAGGACCCCGCGGCTTGGCGTGGAG 654
QY 592 GCGCTGTGAGGACTCGAGTGGTGTGCTGGCGGAGCGCTGGGCGGCGCAACCGCGAG 651
Db 655 GCGCTGTGAGAACACGAGGTGGTGTGCTGGCGTTCAGCGTTCGGGCGGCAACGCTGAG 714
QY 652 GTCATGTACACAGCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGTGAC 711
Db 715 GTCATGTACACAGCTTTCAGGCACTGGCGGACTCCGTTGGACTACGTGCAAGTCTGTGAC 774
QY 712 TCGGACACAAAGTTGGACCCCATGGCACTGCTGGAGCTCTGTCGGGTTACTGGACGAGGAC 771

Db 775 TCAGACACAGACTAGACCCCATGGCACTGCTGGAGCTTGTGCGAGTGTTCGATCAAGAC 834
QY 772 CCCCGGTAGGGCTGTTGGTGGGAGCTGCGGATCCTTAACCTCTGGACTCTCTGGGTC 831
Db 835 CCCCGGTAGGGCTGTTGGGAGGAGTGTGAGGATCCTTAACCTCTGGACTCTCTGGGTC 894
QY 832 AGCTTCTTAAAGCAGCCCTGCGATACTGGGTAGCCCTTCAATGTGGAGCGGCTTGTCAAGAC 891
Db 895 AGCTTCTTAAAGCAGCTTCTGATACTGGGTAGCCCTTCAATGTGGAGCGGCTTGTCAAGAC 954
QY 892 TACTTCCACTGTGTATCTGTCATCAGCGGTTCTCTAGGT 930
Db 955 TACTTCCACTGTGTCTGTCATCAGTGGTCTCTGGGT 993

RESULT 14

ABQ27384
ID ABQ27384 standard; DNA; 662 BP.

XX AC ABQ27384;

XX DT 12-JUL-2002 (first entry)

XX OL Oligonucleotide for detecting cytosine methylation SEQ ID NO 13975.

XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.

XX OS Homo sapiens.

XX PN WO200218632-A2.

XX PD 07-MAR-2002.

XX PF 01-SEP-2001; 2001WO-EP010074.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PR 05-SEP-2000; 2000DE-01044543.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX DR WPI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNPs); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention


```
XX SQ Sequence 662 BP; 80 A; 80 C; 247 G; 255 T; 0 U; 0 Other;
Query Match 39.4%; Score 422; DB 6; Length 662;
Best Local Similarity 77.3%; Pred. No. 1.6e-64;
Matches 512; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
QY 34 GCAGCCCGCCCTGCTCCGGCTGCGCGAGGGTGTGACCATCGCCCTCGCCCTGCTC 93
Db 1 GTAGTTTGTGCTGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTT 60
QY 94 ATCTTGGCCCTCATGACCTGAGCCCTACCGCGCGGGGTGCGCTGCGCTCGATCGCTAC 153
Db 61 ATTTTGGGTTTATGATTTGGGTTTACGTCGTCGGGGTGTGCTTGGTTTTCGATCGTTAC 120
QY 154 GGCCTCTCGCCCTCGCGCTTACCGGGCTTCCCTTTCAGCCACCTGCTGGCGCAGAGC 213
Db 121 GGTATTTTGGTTTGGTTTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 180
QY 214 CTCTTTCGCTACCTGAGCACCAGCGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCG 273
Db 181 TTTTTCGCGTATTTGGAGTATCGCGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGGT 240
QY 274 GCCACCGCGCAGTGTGGCGTACCATCTCCGCTACCGAGGAGCCCGCGTACCTG 333
Db 241 GTTATCGCGCTAGTGTGGCGTGTATTTTTCGTTTATAGGAGGATTTTCGCGTATTTG 300
QY 334 CGCCAGTCTCGCGCTCGCGCGCCCTGCTGTATCCGCGCGCGCGCGCGCGCGCTCCTC 393
Db 301 CGTTAGTGTGGCGTGTGTCGCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTT 360
QY 394 ATGTTGGTGTATGCGAACCGCGCGCGAGGACCTTACATGTCGATGTTCCGCGAGTC 453
Db 361 ATGTTGGTGTATGCGAACCGCGCGAGGATTTTATATGTCGATGTTTTCGCGAGTT 420
QY 454 TTCGCTGACGAGGACCGCGCAGCTACGTTGGGACGCACTACACAGCCCTCGGAA 513
Db 421 TTCGCTGACGAGGATTTTCGTTACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 480
QY 514 CCQCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 573
Db 481 TTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGATTT 540
QY 574 GCGCGCTGCGAGTGTGAGCGCGCTGCTGAGGACTCGCAGGTGCGTGTGCGTGGCGAGCGC 633
Db 541 GCGCGCTTGTGTAGTGTGAGCGCGTGTGTGAGGATTCGTAGTGTGCGTGTGCGTGTGCGT 600
QY 634 TGGGCGCGACGCGAGGTCATGTACAGCCTTCAAGGCGCTCGGAGATTCGTTGGAC 693
Db 601 TGGGCGCGTAAGCGCGAGGTTATGTATATAGTTTAAAGGCGTTCGGAGATTCGTTGGAT 660
QY 694 TA 695
Db 661 TA 662
RESULT 15
ABQ27385/c
ID ABQ27385 standard; DNA; 662 BP.
XX
AC ABQ27385;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 13976.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
```

```
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PF 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
PR 05-SEP-2000; 2000DE-0104543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
DR WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP-B); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ3410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 662 BP; 255 A; 247 C; 80 G; 80 T; 0 U; 0 Other;
Query Match 39.4%; Score 422; DB 6; Length 662;
Best Local Similarity 77.3%; Pred. No. 1.6e-64;
Matches 512; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
QY 34 GCAGCCCGCCCTGCTCCGGCTGCGCGAGGGTGTGACCATCGCCCTCGCCCTGCTC 93
Db 662 GTAGTTTGTGCTGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTT 603
QY 94 ATCTTGGCCCTCATGACCTGAGCCCTACCGCGCGGGGTGCGCTGCGCTCGATCGCTAC 153
Db 602 ATTTTGGGTTTATGATTTGGGTTTACGTCGTCGGGGTGTGCTTGGTTTTCGATCGTTAC 543
QY 154 GGCCTCTCGCCCTTCCGGCTCTACGGGCGCTTCTTTTCAGCGCACCTCGTGGCGAGAGC 213
Db 542 GGTATTTTGGTTTTCGGTTTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCGCGGTTCG 483
QY 214 CTCTTCGCGTACCTGAGACACCGCGGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGATCA 273
Db 482 TTTTTCGCGTATTTGGAGTATCGCGGTGTCGCGCGCGCGCGCGCGCGCGCGCGGTGTA 423
QY 274 GCCACCGCGCAGTGTGGCGCTACCATCTCCGCTACACAGGAGGAGCCCGCGTACCTG 333
Db 422 GTTATCGCGCTAGTGTGGCGTTCGTTTTCGTTTATAGGAGGATTTTCGCGTATTTG 363
QY 334 CGCCAGTGTGCGTGTGCGCGCGCGCGCTGCTGTACCGCGCGCGCGCGCGCGCGCTCCTC 393
Db 362 CGTTAGTGTGTGGCGTTCGTTTCGCGTTCGTTTGTGTTGTTGTTGTTGTTGTTGTTT 303
QY 394 ATGTTGGTGTGATGGCAACCGCGCGCGAGGACCTCTACATGTTGTCGATGTTTCCGCGAGGTC 453
```

[illegible]

Search completed: March 13, 2005, 03:46:24
Job time : 702.034 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 02:13:47 ; Search time 5058.4 Seconds
(without alignments)
10259.279 Million cell updates/sec

Title: US-10-672-399-3

Perfect score: 1071

Sequence: 1 atgagacagcaggacgcgc.....ccggaccatcgctggatga 1071

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.htg.*

3: gb.in.*

4: gb.on.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	927.4	86.6	2088	9 HSU59269	US9269 Human hyalu
2	927.4	86.6	2116	6 AR220003	AR220003 Sequence
3	915.2	85.5	2119	9 BC035837	BC035837 Homo sapi
4	915	85.4	2108	9 HUMHAS	D84424 Homo sapien
5	915	85.4	2117	6 AR137038	AR137038 Sequence
6	854.4	79.8	2109	9 AY463695	AY463695 Papio anu
7	697.2	65.1	229155	9 AC018755	AC018755 Homo sapi
8	679.8	63.5	2095	10 AB097568	AB097568 Rattus no
9	678.2	63.3	2102	6 E13681	E13681 DNA encodin
10	678.2	63.3	2102	6 E28454	E28454 Hyaluronate
11	678.2	63.3	2102	6 E30971	E30971 Hyaluronate
12	678.2	63.3	2102	6 E34326	E34326 DNA for gen
13	678.2	63.3	2102	10 MUSHAS	D82964 Mus musculu
14	669	62.5	193986	2 AC130783	AC130783 Pan trogl
15	635.4	59.3	185623	2 AC137058	AC137058 Papio anu
16	614.8	57.4	1581	6 CQ731079	CQ731079 Sequence
17	492	45.9	293184	2 AC079487	AC079487 Mus muscu
18	490.8	45.8	27955	2 AC108651	Continuation (4 of
19	490.8	45.8	257325	2 AC116203	AC116203 Rattus no

```

c 20 369.2 34.5 611 11 BV099313
c 21 368.2 34.4 611 11 BV160501
c 22 302.4 28.2 1375 5 AY437407
c 23 290.6 27.1 1659 4 AB055979
c 24 274.2 25.6 550 11 BV161377
c 25 272.8 25.5 1662 6 CQ723590
c 26 272.8 25.5 1994 9 AF234839
c 27 272.8 25.5 4049 9 AF232772
c 28 272.8 25.5 4220 6 CQ776471
c 29 265.4 24.8 3995 4 AB159675
c 30 250 23.3 2093 5 BC047963
c 31 248.4 23.2 2950 5 XELDG424
c 32 243.2 22.7 2465 10 AB097569
c 33 240.4 22.4 2234 5 AF015780
c 34 238.8 22.3 2245 5 BC077983
c 35 236.2 22.1 1665 6 E30973
c 36 236.2 22.1 1665 6 AR373269
c 37 236.2 22.1 5919 6 CQ777380
c 38 236.2 22.1 5919 10 MM086408
c 39 207.4 19.4 1229 6 CQ776472
c 40 207.4 19.4 1229 9 BC021853
c 41 195.6 18.3 605 11 BV161384
c 42 194.4 18.2 215158 5 BX539313
c 43 193 18.0 603 11 BV161386
c 44 193 18.0 621 11 BV160507
c 45 192.4 18.0 621 11 BV099322

```

ALIGNMENTS

```

RESULT 1
LOCUS HSU59269 2088 bp mRNA linear PRI 24-SEP-1996
DEFINITION Human hyaluronan synthase mRNA, complete cds.
ACCESSION U59269
VERSION U59269.1 GI:1556464
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2088)
AUTHORS Shyjan,A.M., Heidin,P., Butcher,E.C., Yoshino,T. and Briskin,M.J.
TITLE Functional cloning of the cDNA for a human hyaluronan synthase
J. Biol. Chem. 271 (38), 23395-23399 (1996)
MEDLINE 96394438
PUBMED 8798544
REFERENCE 2 (bases 1 to 2088)
AUTHORS Briskin,M.J. and Shyjan,A.M.
TITLE Direct Submission
JOURNAL Submitted (24-MAY-1996) LeukoSite Inc., 215 First Street,
Cambridge, MA 02142, USA
FEATURES
source
1. .2088
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
36..1772
/codon_start=1
/product="hyaluronan synthase"
/protein_id="AAC50706.1"
/db_xref="GI:1556465"
/translators="MRQDAPKTPAARRCSGLARRVLTAFALLILGLMTWAYAGV
PLASDRYGLLAFLGILGALHVAQSLFAYLEHRRVAAAARGEDFADATASVALTIS
AYQEDPAYLRQCLASARALLYPRARLVLMVDGNRAEDLYMDFMDFEADVPATY
WVDGNYPQWPEPAAGAGAGAVEAEDEGRALVEALVTRCVCAQWGGKREV
MYTAFKALGSDVDYQVCDSDRLDPWALLEVRVLDPRVAGVGDVRLINLDSW
VSFLSLRYVWAFNVACOSYFHCVSCI SPGLGLYRNLLQQFLEAWYNOKFLGTHC
TFGDDRHLTNRMLSMGYATKYTSRSCYSETPSSFLWLSQTRWSKSYFREWLYNAL
WVHKHAWMTTEAVVSLGFFFAVTLRLFYAGRPWALLLVLLCVQVLAQAFAA
WLRGCLRNLLSLYAPLYMCGLLPAKFLALVTMNSGSGSRKRLAANYPLPLAL
WALLLLGLVRSVAHEARDWSGSRAAEAAYHLAAGAGAYVGYVWVWVRL

```


Db 756 AGGTTGGACCCCATGCGCACTGCTGGAGCTGTGCGGGTACTGGACGAGACCCCGGGTA 815
 Qy GGGGCTGTTGGTGGGACGTCGGGATCCTTAACCCCTCTGGACTCTCGGTGAGCTTCTCTA 840
 Db GGGGCTGTTGGTGGGACGTCGGGATCCTTAACCCCTCTGGACTCTCGGTGAGCTTCTCTA 875
 Qy AGCAGCTGCGATACCTGCGGTAGCTTCAATGTGGAGCGGGCTTGTGACAGCTTCTCCAC 900
 Db AGCAGCTGCGATACCTGCGGTAGCTTCAATGTGGAGCGGGCTTGTGACAGCTTCTCCAC 935
 Qy TGTGTATCTCTGATCAGCGGTCTCTAGG 929
 Db TGTGTATCTCTGATCAGCGGTCTCTAGG 964

RESULT 3
 BC035837
 LOCUS BC035837 2119 bp mRNA linear PRI 30-JUN-2004
 DEFINITION Homo sapiens hyaluronan synthase 1, mRNA (cDNA clone MGC:46218
 IMAGE:5589083), complete cds.

ACCESSION BC035837
 VERSION BC035837.1 GI:23243101

KEYWORDS MGC.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullah, S.J., Bosak, S.A., McSwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 2119)

Strausberg, R.

Direct Submission

Submitted (31-JUL-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Invitrogen

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc_mgc@hgrl.nih.gov

Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.L., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAP Plate: 79 Row: f Column: 24

This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 4504338.

Location/Qualifiers

1. 2119

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:46218 IMAGE:5589083"

/tissue_type="Ovary, pooled from 3 adults"

/clone_lib="NIH MGC_125"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. 2119

/gene="HAS1"

/note="synonym: HAS"

/db_xref="LocusID:3036"

/db_xref="MIM:601463"

42. 1775

/gene="HAS1"

/codon_start=1

/product="hyaluronan synthase 1"

/protein_id="AAH35837.1"

/db_xref="GI:23243102"

/db_xref="LocusID:3036"

/db_xref="MIM:601463"

/translation="MRQDAKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVP
 LASPYGLLAGLYGALSAHLVAQSLPAYLEHRRVAAARGPDLDAATARSVALTISA
 YQEDPAYRQCLASARALLPRARLRLVLMVVDGNRAEDLYWDMFREVFAEDPATYV
 WDGNHQPWEPAAGAGAGAYREVAEDPGRLEAVEALVRRRCVCAARGGKREVM
 YTFKALGDSMDYVQVCDSTRDLPMALLELRVLDDEPRVAGVGGDVRILNPLDSW
 FGLSLRYWFMNVERACQSYFHCVCYSETSPSLRWLSQCTRRKSKSVFRWLYNALW
 WRRHAWTYEAVVSGLPFPEFAVATLRLFYAGRPWALLVLLCVQVALAKAAPAAW
 LRGLRWLISLYAPLWCGLLPAKFLALVTMNSGWTSGRRKLAAAYVPLPLALW
 ALLLGGILRSVAHEARADWSGPRAAEAYHLLAAGAYGVYVWAMLTLYWVVRRLC
 RRRTGGYRVQV"

ORIGIN

Query Match 85.5%; Score 915.2; DB 9; Length 2119;

Best Local Similarity 99.7%; Pred. No. 4.5e-112;

Matches 917; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 10 CAGGACGCGCCCAAGCCACCTCTCGACGCCCGCGCTGCTCGGCGCTGGCCCGGAGGGTG 69

Db 48 CAGGACGCGCCCAAGCCACCTCTCGACGCCCGCGCTGCTCGGCGCTGGCCCGGAGGGTG 107

Qy 70 CTGACCATCGCTTGGCGCTGCTCATCTGGCGCTCATGACCTGGCGCTACGCGCCGG 129

Db 108 CTGACCATCGCTTGGCGCTGCTCATCTGGCGCTCATGACCTGGCGCTACGCGCCGG 167

Qy 130 GTGCGCTGGCGCTCGATCGCTAGCGCTCTCGGCGCTTCGGCGCTTCTTCCTT 189

Db 168 GTGCGCTGGCGCTCGATCGCTAGCGCTCTCGGCGCTTCTTCCTT 227

Qy 190 TCAGCGCACCTGTGGCGCAGAGCTCTTTCGCTACCTGGAGCACCGCGCGGTGGCGCG 249

Db 228 TCAGCGCACCTGTGGCGCAGAGCTCTTTCGCTACCTGGAGCACCGCGCGGTGGCGCG 287

Qy 250 GCGCGCGCGCGCGCTGGATGACGCCACCGCGCGCAGGTGTGGCGCTGACCATCTCCGCC 309

Db 288 GCGCGCGCGCGCGCTGGATGACGCCACCGCGCGCAGGTGTGGCGCTGACCATCTCCGCC 347

Qy 310 TACAGAGGAGACCGCGGTACTCTGGCGCAGTGCCTTGGCGCTCGCGCGCGCTTGTGTAC 369

Db 348 TACAGAGGAGACCGCGGTACTCTGGCGCAGTGCCTTGGCGCTCGCGCGCGCTTGTGTAC 407

```
Qy 370 CCGCGCGCGCGTGGCGTCTCATGTGTGTGATGGCAACCGCGCGGAGACTCTAC 429
Db |||||
Qy 408 CCGCGCGCGCGTGGCGTCTCATGTGTGTGATGGCAACCGCGCGGAGACTCTAC 467
Db |||||
Qy 430 ATGCTCGACATGTTCCGCGAGGTCTTCCTGCTGACGAGGACCCCGCCACATGCTGTGGAC 489
Db |||||
Qy 468 ATGCTCGACATGTTCCGCGAGGTCTTCCTGCTGACGAGGACCCCGCCACATGCTGTGGAC 527
Qy 490 GGCACACTACCAACGACCTTGGGAACCCCGCGCGGCGGCGGTGGCGCGGAGCCTAT 549
Db |||||
Qy 528 GGCACACTACCAACGACCTTGGGAACCCCGCGCGGCGGCGGTGGCGCGGAGCCTAT 587
Qy 550 CGGAGGTGGAGGCGGAGGATCTTGGGCGGCTGGCAGTGGAGGCGCTGGTGGAGACTCGC 609
Db |||||
Qy 588 CGGAGGTGGAGGCGGAGGATCTTGGGCGGCTGGCAGTGGAGGCGCTGGTGGAGACTCGT 647
Qy 610 AGGTGCGTGTGCTGGCGCAGCGCTGGGCGCGGCAAGCGCGAGGTCTGCTGACAGCCTTC 669
Db |||||
Qy 648 AGGTGCGTGTGCTGGCGCAGCGCTGGGCGCGGCAAGCGCGAGGTCTGCTGACAGCCTTC 707
Qy 670 AAGCGCTCGAGATGCTGGTGGACTACGTGAGGTCTGTGACTCGGACACAGGTTGGAC 729
Db |||||
Qy 708 AAGCGCTCGAGATGCTGGTGGACTACGTGAGGTCTGTGACTCGGACACAGGTTGGAC 767
Qy 730 CCCATGGCACGTGGAGCTGTGCGGTACTGGACGAGGACCCCGGTTAGGGCTGTT 789
Db |||||
Qy 768 CCCATGGCACGTGGAGCTGTGCGGTACTGGACGAGGACCCCGGTTAGGGCTGTT 827
Qy 790 GGTGGGACGTGGGATCTTAACCCCTCTGACTCCTGGGTGAGTCTTAAGCAGCCTG 849
Db |||||
Qy 828 GGTGGGACGTGGGATCTTAACCCCTCTGACTCCTGGGTGAGTCTTAAGCAGCCTG 887
Qy 850 CGATCTGGGTAGCTTCAATGTGGAGCGGCGCTGTGAGCTACTTCCACTGTGTATCC 909
Db |||||
Qy 888 CGATCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGCTACTTCCACTGTGTATCC 947
Qy 910 TGCATCAGCGTCTCTAGG 929
Db |||||
Qy 948 TGCATCAGCGTCTCTAGG 967
```

RESULT 4
LOCUS HUMHAS 2108 bp mRNA linear PRI 06-FEB-1999
DEFINITION Homo sapiens mRNA for hyaluronan synthase, complete cds.

ACCESSION D84424
VERSION D84424.1 GI:1401033
KEYWORDS hHAS; hyaluronan synthase.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 2108)
AUTHORS Itano,N. and Kimata,K.
TITLE Molecular cloning of human hyaluronan synthase
JOURNAL Blochem. Biophys. Res. Commun. 222 (3), 816-820 (1996)
MEDLINE 96244584
PUBMED 8651928

REFERENCE 2 (bases 1 to 2108)
AUTHORS Itano,N.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2108)
AUTHORS Itano,N.

Direct Submission
TITLE Submitted (16-APR-1996) Naoki Itano, Aichi Medical University,
Institute for Molecular Science of Medicine, Nagakute, Aichi-gun,
Aichi 480-11, Japan (Tel:052-264-4811(ex.2087), Fax:0561-63-3532)

FEATURES
Location/Qualifiers
1..2108
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="brain"

```
gene  
/clone_lib="/lg11"  
/dev_stage="fetal"  
1..2108  
/gene="hHAS"  
149..1780  
/gene="hHAS"  
/codon_start=1  
/product="hyaluronan synthase"  
/protein_id="BAAL2351.1"  
/db_xref="GI:1401033"  
/translation="MTWAYAAGVPLASDRYGLLAGFLYGAFLSHLVAQSLFAVLEHR  
RVAAARGPLDAATARSVALTI SAYQDNTPAYLRQCLASARALLYPRARVRLVMVDGN  
RADLXYMDNREVFADEDPATYVMDNTHQPMWPAAGAGVAGAYREVAREDEPRLA  
VEALVTRRCVCAVRWGGKREVMYTAFAKALGSDVDYVQVCDSDTRDPMALLBLVRV  
LDEDPVGAAGGVDVRLINPLDSWVSLSSLYMVAFNVERACOSYFVCSICSGPLGI  
YRNLIQFLFLEAWYKFLGTFGDDRLHTRNLSMGATYATKTSRSRCVSETPSPF  
LRWLSQOTRWSKSYFREWLYNALMHRHAWMTYEAUVGLFFPFAVFLVRLPYAGR  
PALLWVLLCVQGVAKAAFAAALWRLCLRMVLLSLYAPLMCGLLPAKFLALVTMQ  
SGWTSRRKLAANYVPLPLALWALLLGLVRSVAHEARADWSGSPRAAEAYHLAA  
GAGAYGVYVWAMLTLYVGVVRLCRRTGGRVQV"
```

ORIGIN

```
Query Match 85.4%; Score 915; DB 9; Length 2108;  
Best Local Similarity 99.5%; Pred. No. 4.8e-112;  
Matches 918; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
Qy 7 CAGCAGACGCGCGCCCAAGCCCACTCTCTGACGCCCGCGCTCTCCGCCCTGGCCGCGGAGG 66  
Db |||||  
Qy 50 CAGCAGACGCGCGCCCAAGCCCACTCTCTGACGCCCGCGCTCTCCGCCCTGGCCGCGGAGG 109  
Qy 67 GTGCTGACCATCGGCTTTCGCCCTCATCTCTGGGCTCATGACCTGGGCTACGCCCGCC 126  
Db |||||  
Qy 110 GTGCTGACCATCGGCTTTCGCCCTCATCTCTGGGCTCATGACCTGGGCTACGCCCGCC 169  
Qy 127 GGGTGGCGTGGCTCCGATCGCTACGGCTCTCTGGCTTCGGCTCTACGGGCGCTTC 186  
Db |||||  
Qy 170 GGGTGGCGTGGCTCCGATCGCTACGGCTCTCTGGCTTCGGCTCTACGGGCGCTTC 229  
Qy 187 CTTCAGCGCACCTGGTGGCGGAGAGCTCTTCGGCTACCTGGGACACCGCGCGGTGGCG 246  
Db |||||  
Qy 230 CTTCAGCGCACCTGGTGGCGGAGAGCTCTTCGGCTACCTGGGACACCGCGCGGTGGCG 289  
Qy 247 GCGCGCGCGCGCGCGCTGGATGACCCACCGCGCGAGTGTGGCGCTGACCATCTCC 306  
Db |||||  
Qy 290 GCGCGCGCGCGCGCGCTGGATGACCCACCGCGCGAGTGTGGCGCTGACCATCTCC 349  
Qy 307 GCTTACAGAGAGACCCCGCTGCTGCGGCTCTGGCTTCGGCTTCGGCGCGCGCTTCGTG 366  
Db |||||  
Qy 350 GCTTACAGAGAGACCCCGCTGCTGCGGCTACCTGCGGCTGCGCTTCGGCGCGCGCTTCGTG 409  
Qy 367 TACCGCGCGCGCGCTGCGGCTCTCATGCTGGTGGATGGACACCGCGCGGAGACCTC 426  
Db |||||  
Qy 410 TACCGCGCGCGCGCTGCGGCTCTCATGCTGGTGGATGGACACCGCGCGGAGACCTC 469  
Qy 427 TACATGTTCACATGTTCCCGGAGGTCTTCGCTGACGAGGACCCCGCGCTGCTGG 486  
Db |||||  
Qy 470 TACATGTTCACATGTTCCCGGAGGTCTTCGCTGACGAGGACCCCGCGCTGCTGG 529  
Qy 487 GACGCGCAACTACCAACGACCTTGGGAACCCCGCGCGCGCGCTGGGCGCGCGGAGCC 546  
Db |||||  
Qy 530 GACGCGCAACTACCAACGACCTTGGGAACCCCGCGCGCGCGCTGGGCGCGCGGAGCC 589  
Qy 547 TATCGGAGGTGGAGCGGAGGATCTTCGGCGGCTGGCAGTGGAGGCGCTGGTGGAGACT 606  
Db |||||  
Qy 590 TATCGGAGGTGGAGCGGAGGATCTTCGGCGGCTGGCAGTGGAGGCGCTGGTGGAGACT 649  
Qy 607 CGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGCGCGCAAGCGCGGAGGTCTGTACACAGCC 666  
Db |||||  
Qy 650 CGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGCGCGCAAGCGCGGAGGTCTGTACACAGCC 709  
Qy 667 TTCAGGCGCTCGGAGATTTCGGTGGACTACGTGAGGTCTGTGACTCGGACACAGGTTG 726  
Db |||||  
Qy 710 TTCAGGCGCTCGGAGATTTCGGTGGACTACGTGAGGTCTGTGACTCGGACACAGGTTG 769
```

```

QY 727 GACCCCATGGCACTCTGAGCTCTGCGGGTACTGGACGAGGACCCCGGGTAGGGCT 786
Db 770 GACCCCATGGCACTCTGAGCTCTGCGGGTACTGGACGAGGACCCCGGGTAGGGCT 829
QY 787 GTTGGTGGGAGCTGGGATCCTTAACCTCTGGACTCTGGTCACTTCTTAAGCAGC 846
Db 830 GTTGGTGGGAGCTGGGATCCTTAACCTCTGGACTCTGGTCACTTCTTAAGCAGC 889
QY 847 CTGCGATCTGGGTAGCTTCAATCTGGAGCGGGCTTGTCAAGCTACTTCCACTGTGA 906
Db 890 CTGCGATCTGGGTAGCTTCAATCTGGAGCGGGCTTGTCAAGCTACTTCCACTGTGA 949
QY 907 TCCTGCATCAGCGGTCTCTAGG 929
Db 950 TCCTGCATCAGCGGTCTCTAGG 972

RESULT 5
ARI37038
LOCUS ARI37038 2117 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6162908.
ACCESSION ARI37038
VERSION ARI37038.1 GI:14478288
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE
AUTHORS Itano,N. and Kimata,K.
TITLE Polypeptide of human-origin hyaluronate synthetase and DNA encoding the same
JOURNAL Patent: US 6162908-A 1 19-DEC-2000;
FEATURES
Source Location/Qualifiers
1..2117
/mol_type="unassigned DNA"
/organism="unknown"

ORIGIN
Query Match 85.4%; Score 915; DB 6; Length 2117;
Best Local Similarity 99.5%; Pred. No. 4,8e-112;
Matches 918; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 7 CAGCAGGACGCGCCCAAGCCCACTCTGCAGCGCCGCTGCTCGGCTGCGCCGAGG 66
Db 50 CAGCAGGACGCGCCCAAGCCCACTCTGCAGCGCCGCTGCTCGGCTGCGCCGAGG 109
QY 67 GTGCTGACATCGCTTGGCTGCTATCTGGGCTCATGACCTGGGCTTACCGGCC 126
Db 110 GTGCTGACATCGCTTGGCTGCTATCTGGGCTCATGACCTGGGCTTACCGGCC 169
QY 127 GGGGTGCGCTGGCTCGATCGCTACGCGCTCTGCGCTTGGCTTACGCGGCTTC 186
Db 170 GGGGTGCGCTGGCTCGATCGCTACGCGCTCTGCGCTTGGCTTACGCGGCTTC 229
QY 187 CTTTCAGCGCACCTGGTGGCGCAGAGCTCTTTCGGCTACCTGGAGCACCGCGGGTGCG 246
Db 230 CTTTCAGCGCACCTGGTGGCGCAGAGCTCTTTCGGCTACCTGGAGCACCGCGGGTGCG 289
QY 247 CGCGCGCGCGGGGCGCTGTGATGACGACACCGCGCGAGTGTGGCGCTGACCATCTCC 306
Db 290 CGCGCGCGCGGGGCGCTGTGATGACGACACCGCGCGAGTGTGGCGCTGACCATCTCC 349
QY 307 GCCTTACAGGAGGACCCCGCTGCTGCTGCGCGAGTGTGGCTGCGCGCGCGCTGCTG 366
Db 350 GCCTTACAGGAGGACCCCGCTGCTGCTGCGCGAGTGTGGCTGCGCGCGCGCTGCTG 409
QY 367 TACCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 426
Db 410 TACCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 469
QY 427 TACATGTCGATGTTTCGCGAGGCTTTCGCTGACGAGGACCCCGCCACGTCAGTGTGG 486

```

```

Db 470 TACATGTCGATGTTTCGCGAGGCTTTCGCTGACGAGGACCCCGCCACGTCAGTGTGG 529
QY 487 GACGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 546
Db 530 GACGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 589
QY 547 TATCGGAGGTGGAGGCGGAGATCTCTGGCGGGCTGGCAGTGGAGCGCTGGTGAAGACT 606
Db 590 TATCGGAGGTGGAGGCGGAGATCTCTGGCGGGCTGGCAGTGGAGCGCTGGTGAAGACT 649
QY 607 CGCAGGTGCGTGGTGGCGCGAGGCTGGGGCGGCAAGCGGAGGTCTATGTACACAGCC 666
Db 650 CGCAGGTGCGTGGTGGCGCGAGGCTGGGGCGGCAAGCGGAGGTCTATGTACACAGCC 709
QY 667 TTCAAGGCGCTCGGAGATTCGCTGAGTACGTGACGCTCTGACTCCGACACAAAGGTTG 726
Db 710 TTCAAGGCGCTCGGAGATTCGCTGAGTACGTGACGCTCTGACTCCGACACAAAGGTTG 769
QY 727 GACCCCATGGCACTCTGCGGCTCTGCGGCTTCAATCTGGAGCGGCTTTCAGAGCTACTTCCACTGTGA 786
Db 770 GACCCCATGGCACTCTGCGGCTCTGCGGCTTCAATCTGGAGCGGCTTTCAGAGCTACTTCCACTGTGA 829
QY 787 GTTGGTGGGAGCTGGGATCCTTAACCTCTGGACTCTGGTCACTTCTTAAGCAGC 846
Db 830 GTTGGTGGGAGCTGGGATCCTTAACCTCTGGACTCTGGTCACTTCTTAAGCAGC 889
QY 847 CTGCGATCTGGGTAGCTTCAATCTGGAGCGGGCTTGTCAAGCTACTTCCACTGTGA 906
Db 890 CTGCGATCTGGGTAGCTTCAATCTGGAGCGGGCTTGTCAAGCTACTTCCACTGTGA 949
QY 907 TCCTGCATCAGCGGTCTCTAGG 929
Db 950 TCCTGCATCAGCGGTCTCTAGG 972

RESULT 6
AY463695
LOCUS AY463695 2109 bp mRNA linear PRI 07-DEC-2003
DEFINITION Papio anubis hyaluronan synthase (HAS1) mRNA, complete cds.
ACCESSION AY463695
VERSION AY463695.1 GI:38607341
KEYWORDS
SOURCE Papio anubis (olive baboon)
ORGANISM Papio anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopitheciinae; Papio.
REFERENCE
AUTHORS Martinez-Duncker,I., Oriol,R. and Mollicone,R.
TITLE Evolution of the the hyaluronan, nodulation c, chitin and cellulose synthases: a superfamily of cell-wall associated carbohydrate polymerizing enzymes
JOURNAL Unpublished
REFERENCE
AUTHORS Martinez-Duncker,I., Oriol,R. and Mollicone,R.
TITLE Direct Submission
JOURNAL Submitted (12-NOV-2003) Unite 504 Glycobiology Et Signalisation Cellulaire, Inserm, 16 Avenue Paul Vaillant Couturier, Villejuif 94807, France
FEATURES
source Location/Qualifiers
1..2109
/organism="Papio anubis"
/mol_type="mRNA"
/db_xref="taxon:9555"
1..2109
/gene="HAS1"
36..1787
/gene="HAS1"
/note="cell wall associated protein; produces hyaluronic acid chains"
/codon_start=1
/product="hyaluronan synthase"
/protein_id="AAR25554.1"

```



```
/db_xref="GI:38607342"
/translation="MTORDTPKTPPAARCSGLARRVLTIAPALLILGLMTWAAGV
PLASDRYGLIAFGYGLPLAHLAQSLFAYLEHRRVAAARPAARAAARLDAATARSV
ALTTISAYOEDPAYLRQCLVLSARALLYPRALRLVLMVDGNRPEDALRYMDFREYFADE
DPATYVWDGNYPHQWPEAAGVAGAYREVEADPGRLAVALYRTRCVCVQAORWG
GKREVMYTAFAKALGDSVDYVQVCDSDTRLPDMALLELVQVLDEDPRVGAGGVDIRILN
PLDSWVSFLASLRVYVAFNVERRACQSVPHVCISLGPLGLRYNNLLQOFLAEWYNOKP
LGHTCTGDDRHLLTNMLSMGYATKYSRSRCYSETPSSFLRWLSQOTRWSKSYFREM
LYNALWHRHWMWYTAENVGLFFPFFVATATVLRFLYAGRPWALLWVLLCVQGVALAK
AAFAALURGLCLRMVLISLYAPLYMCGLLPAKFLALVTMNSQMGWTSRRRLKLAANYPL
LPLALWALLGLLGLSLVSHVHEARADWSPSRAAEYHLAAGAGYVGVVWVMTLYWV
GVRRLCRRTGGYRVQV"
```

ORIGIN	
Query Match	79.8%; Score 854.4; DB 9; Length 2109;
Best Local Similarity	95.1%; Pred. No. 5e-104;
Matches 898; Conservative 0; Mismatches 31; Indels 15; Gaps 1;	
Qy	1 ATGAGACAGCAGGACGCGCCCAAGCCACTCTCGAGCCGCGCGCTGCTCGGCGCTCGCC 60
Db	
Qy	36 ATGACACAGCGGAGACGCGCCCAAGCCACTCTCGAGCCGCGCGCTGCTCGGCGCTCGCC 95
Db	
Qy	61 CGGAGGCTGTGACCATCGCCCTTCGCCCTGTCTATCTCTGGGCTCATGACCTGGGCGCTAC 120
Db	
Qy	121 GCCCGCGGGTGCCTGGGCTTCGATCGCTACGGCTCTCTGGGCTTCGGGCTCTACGGG 180
Db	
Qy	156 GCCCGCGGGTGCCTGGGCTTCGATCGCTACGGCTCTCTGGGCTTCGGGCTCTACGGG 215
Db	
Qy	181 GCCTTCCTTTTCAGCGCACCTGCTGGCGCAGAGCTCTTCGGGTACCTGGAGCACCGGCGG 240
Db	
Qy	216 GCCTTCCTCTCGGCGCACCTGTTGGCGAGAGCTCTTCGGGTACCTGGAGCATCGGCGG 275
Db	
Qy	241 GTGCGCGCGCGCGG-----CGGCGCGCGCTGTGATGACGACCGACCGCGCGC 285
Db	
Qy	276 GTGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGTGATGACGACCGACCGCGCGC 335
Db	
Qy	286 AGTGTGGCGGTGACCATCTTCGCCCTACAGAGAGACCGCGGTACCTGGCGCAGTGCGCTG 345
Db	
Qy	336 AGCGTGGCGGTGACCATTTTCGCCCTACAGAGAGACCGCGGTACCTGGCGCAGTGCGCTG 395
Db	
Qy	346 GCGTCCGCGCGCGCTGTCTACCGCGCGCGGTGCGCGTCTCATGCTGGTGGTGGAT 405
Db	
Qy	396 GTGTCCGCGCGCGCTGTCTGTATCCGCGCGCGCGCTGCGCGTCTCATGCTGGTGGAC 455
Db	
Qy	406 GGCAACCGCGCGGAGGACCTCTACATGGTGCACATGTTTCGCGAGGCTTTCGCTGACGAG 465
Db	
Qy	456 GGCAACCGCGCGGAGGACCTCTACATGGTGCACATGTTTCGCGAGGCTTTCGCGCAGAG 515
Db	
Qy	466 GACCCCGCAGTACGTGTGGAGCGGCACTACACACGCGCTGGGAACCGCGCGCGCGG 525
Db	
Qy	516 GACCCCGCAGTACGTGTGGAGCGGCACTACACACGCGCTGGGAACCGCGCGCGGTG 575
Db	
Qy	526 GGCGCGTGGCGCGCGAGGCTATCGGAGGTGAGCGGAGGATCTCGGCGCGCTCGCA 585
Db	
Qy	576 GGCGCGTGGCGGTGCGAGGCTTACCGGAGGTGAGCGGAGGATCCGCGCGGTGTGCG 635
Db	
Qy	586 GTGAGGCGCTGTGAGGACTCGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGGCGGAAG 645
Db	
Qy	636 GTGAGGCGCTGTGAGGACTCGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGGCGGCAAG 695
Db	
Qy	646 CGCAGGTGATGTACACAGCTTACAGCGCTCGAGATTCGCTGGAGCTACGTGCAGGTC 705
Db	
Qy	696 CGCAGGTGATGTACACAGCTTACAGCGCTCGAGACTCGGTGGAGCTACGTGCAGGTC 755
Db	
Qy	706 TGTGACTCGGACACAAAGTTGAGCCCATGCTGCTGGAGCTCGTGGGCTACTGGAC 765
Db	
Qy	756 TGTGACTCGGACACAAAGTTGAGCCCATGCTGCTGGAGCTCGTGGAGTCTCTGGAT 815
Db	
Qy	766 GAGGACCCCGGGTAGGGGCTGTTGGTGGGACGTGGGATCTTAAACCTCTGGACTCC 825
Db	
Qy	816 GAGGACCCCGGGTAGGGGCTGTTGGTGGGACGTGGGATCTTAAACCTCTGGACTCC 875
Db	

Qy	826 TGGGTACAGCTTCTTAAGCAGCTGCGATACCTGGGTAGCTTCAATGTGGAGCGGCTTGT 885
Db	
Qy	876 TGGGTACAGCTTCTTAAGCAGCTGCGATACCTGGGTAGCTTCAATGTGGAGCGGCTTGT 935
Db	
Qy	886 CAGAGCTACTTCCACTGTGTATCTGTCATCAGCGGTTCTCTAGG 929
Db	
Qy	936 CAGAGCTACTTCCACTGTGTCTCTGTCATCAGTGTCTCTAGG 979
Db	
RESULT 7	
AC018755/c	229155 bp DNA linear PRI 26-JUL-2000
LOCUS	Homo sapiens chromosome 19, BAC BC330783 (CIT-HSPC_470E3), complete
DEFINITION	Homo sapiens sequence.
ACCESSION	AC018755
VERSION	AC018755.3 GI:9454515
KEYWORDS	HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	1 (bases 1 to 229155)
AUTHORS	Lamerdin,J.E., McCreedy,P.M., Skowronski,E., Viswanathan,V., Burkhardt-Schultz,K., Gordon,L., Diae,J., Ramirez,M., Stilwagen,S., Phan,H., Velasco,N., Do,L., Regala,W., Terry,A., Brower,A., Gaines,J., Danganan,L., Erler,A., Christensen,M., Georgescu,A., Avila,J., Liu,S., Andreise,T., Frankheim,M., Attix,C., Anico-Keller,G., Coefield,J., Duarte,S., Lucas,S., Bruce,R., Thomas,P., Quan,G., Kronmiller,B., Arellano,A., Sanders,C., Ow,D., Nolan,M., Trong,S., Kobayashi,A., Olsen,A.S. and Carrano,A.V.
TITLE	Sequence analysis of a 5-Mb region in 19q13.4
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 229155)
AUTHORS	DOE Joint Genome Institute.
TITLE	Direct Submission
JOURNAL	Submitted (19-DEC-1999) Production Sequencing Facility, DOE Joint Genome Institute. 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE	3 (bases 1 to 229155)
AUTHORS	Lamerdin,J.E.
TITLE	Direct Submission
JOURNAL	Submitted (26-JUL-2000) DOE Joint Genome Institute, Lawrence Livermore National Laboratory, Livermore, CA 94550
COMMENT	On Jul 26, 2000 this sequence version replaced gi:7458780. Map and sequence oriented from centromere to q-telomere. BC330783 (CIT-HSPC_470E3) is overlapped on the left by BC849408 and overlaps cosmid R28782 (LLNL-R_248F10, AC005946) on the right from bases 217,905 to 229,155 of this accession. Additional chromosome 19 map and sequence information may be obtained at: http://www-bio.lnl.gov/bbrp/genome/genome.html .
FEATURES	Location/Qualifiers
source	1. .229155
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="19"
	/map="19q13.4"
	/clone="CITC-470E3 (LLNL clone name BC330783)"
	/clone_lib="Caltech human BAC library C"
repeat_region	201..262
	/rpt_family="L2"
gene	<299..16870
	/gene="ZNF175"
mRNA	/note="zinc finger protein 175; synonym: OTK18"
	join(<299..550,8540..8666,8936..9031,13776..16870)
	/gene="ZNF175"
	/product="ZNF175 mRNA"
CDS	join(479..550,8540..8666,8936..9031,13776..15616)
	/gene="ZNF175"
	/codon_start=1
	/product="zinc finger protein 175"
	/protein_id="AAF87843.1"

```
/db_xref="GI:9454517"
/translation="MPADVNLISQRPQVLGPKEQDSCBASVSFEDVTVDPSREBQQOL
DPAQCLRYDVMLELYSHLFAVGTHIPNPEVIFMLKEKEPRVEAEVSHQRCOREP
GLEIPQKEITSKCSQFQKMGKIRTRPHLASSQKOPKCCLPFESI.KLNLFVNGONESN
PFNKTLTNSCNSQKOPKMGKIRTRPHLASSQKOPKCCLPFESI.KLNLFVNGONESN
DTEQLDDVVGSGQLFSHSSDACSNNHTGTFTFCGNOCRAVCGHKQSLKQHQHTQK
KPDGCEGGSFTQKSHLFAQORLHSVGNLHCCGCGKAFQMLSLVILYDHTGDIP
CICKEGKVTFORSELTHQKTRKPKYKDCGKAFQMLSLFRHQRTHSREKLYE
CEGCGKFSQNSTLIHQKHTGTGRQVACSGKAFQKSTLSLHQRTHSQKQSVCCI
ECGQAFQKAHLIWHQSHTGKPKYQCHNCGKFSISKOLDIHHRIHTGKPKYCSDC
KKTQKSHLNIHOKHTGERHHVCEGCGKAFNOKSILSMHORLHTGKPKYKCEGK
AFTSKSQPKHORLHTGKPKYVTECGKAFNRSNPHKHQIHTHRRFPVYKCGKAF
VQKSELTHQRIHMERKEPYECDCGKFSKAPQLKQRIHTGERPYVCEGCGKAFNN
RSNENKQHTTHTRDKSKYCSYKVGFTKQ"
repeat_region
1097..1198
/rpt_family="MER99"
repeat_region
1559..1829
/rpt_family="AluJb"
/rpt_family="AluJb"
complement(2130..2355)
/rpt_family="LIME"
repeat_region
complement(2359..2627)
/rpt_family="AluY"
repeat_region
complement(2816..3183)
/rpt_family="HAL1"
repeat_region
complement(3290..3598)
/rpt_family="AluSg"
repeat_region
complement(3611..3760)
/rpt_family="MIR"
repeat_region
complement(3766..4067)
/rpt_family="HAL1"
repeat_region
4098..4215
/rpt_family="AluJb"
repeat_region
complement(4339..4406)
/rpt_family="HAL1"
repeat_region
complement(4518..4724)
/rpt_family="HAL1"
repeat_region
complement(4803..5110)
/rpt_family="AluY"
repeat_region
5266..5561
/rpt_family="AluSg"
repeat_region
5635..5932
/rpt_family="AluJb"
repeat_region
5984..6120
/rpt_family="AluJb"
repeat_region
6121..6423
/rpt_family="AluSx"
repeat_region
6424..6579
/rpt_family="AluJb"
repeat_region
6647..6913
/rpt_family="LIME"
repeat_region
6936..7295
/rpt_family="MER33"
repeat_region
7700..7785
/rpt_family="L2"
repeat_region
complement(9216..9434)
/rpt_family="LIME"
repeat_region
complement(9435..9745)
/rpt_family="AluY"
repeat_region
complement(9746..9815)
/rpt_family="LIME"
repeat_region
complement(10009..10303)
/rpt_family="AluSp"
repeat_region
11783..12637
/rpt_family="L1MB2"
STS
12180..12330
/gene="ZNFI75"
/standard_name="SHGC-32377"
repeat_region
13436..13619
/rpt_family="MIR"
repeat_region
16068..16112
/rpt_family="(TA)n"
repeat_region
complement(16113..16329)
/rpt_family="MER46A"
```

```
repeat_region
16330..16390
/rpt_family="(TA)n"
repeat_region
16403..16439
/rpt_family="(TATG)n"
repeat_region
16793..17013
/rpt_family="MER58A"
repeat_region
17024..17052
/rpt_family="(CA)n"
repeat_region
17311..17545
/rpt_family="MIR"
complement(18683..18896)
/rpt_family="MIR"
repeat_region
complement(18915..19183)
/rpt_family="AluY"
repeat_region
19336..19599
/rpt_family="MIR"
repeat_region
19811..19928
/rpt_family="CA-rich"
repeat_region
22181..23081
/rpt_family="L2"
repeat_region
complement(23692..23901)
/rpt_family="L2"
repeat_region
complement(23902..24201)
/rpt_family="AluSg1"
repeat_region
complement(24202..24651)
/rpt_family="L2"
repeat_region
26136..26434
/rpt_family="LTR12"
repeat_region
26435..26525
/rpt_family="LTR30"
repeat_region
26526..26643
/rpt_family="LTR12"
repeat_region
complement(26686..26884)
/rpt_family="AluSg/x"
repeat_region
27065..27240
/rpt_family="MIR"
repeat_region
complement(27238..27554)
/rpt_family="L2"
repeat_region
28406..28433
/rpt_family="GC-rich"
repeat_region
28641..28678
/rpt_family="GC-rich"
repeat_region
complement(28997..29259)
/rpt_family="MSR1"
repeat_region
complement(30219..30489)
/rpt_family="L2"
repeat_region
30753..30882
/rpt_family="FLAM_C"
repeat_region
31138..31416
/rpt_family="AluY"
repeat_region
complement(31547..31716)
/rpt_family="MER3"
repeat_region
32041..32112
/rpt_family="LIMEc"
repeat_region
32117..32516
/rpt_family="MLT2E"
repeat_region
35056..35118
/rpt_family="MIR"
repeat_region
complement(35578..35636)
/rpt_family="MERSA"
repeat_region
complement(35690..35820)
/rpt_family="MIR"
repeat_region
35872..35902
/rpt_family="(TG)n"
repeat_region
complement(35906..36209)
/rpt_family="AluJo"
```

Query Match 65.1%; Score 697.2; DB 9; Length 229155;
Best Local Similarity 99.6%; Pred. No. 8.6e-84;
Matches 699; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 7 CAGCAGACGCGCCCAAGCCCACTCTCGACCGCGCGCTCTCGGCTGCGCCGAGG 66
|||||

Db 147050 CAGCAGGAGCGCCCAAGCCACCTCTGACGCTGCGCGCTCTCGCGGAGG 146991
Qy 67 GTGCTGACCATCGCCTTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCCGCC 126
Db 146990 GTGCTGACCATCGCCTTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCCGCC 146931
Qy 127 GGGGTGCGCGCTGCGCTCGCATGCTGACGGCTCTCTGGGCTTTCGCCCTTACGGGCGCTTTC 186
Db 146930 GGGGTGCGCGCTGCGCTCGCATGCTGACGGCTCTCTGGGCTTTCGCCCTTACGGGCGCTTTC 146871
Qy 187 CTTTCAGCGCACCTGCTGCGCGCAGAGCCTCTTCGCGTACCTGAGACCGCGCGGCTGCGC 246
Db 146870 CTTTCAGCGCACCTGCTGCGCGCAGAGCCTCTTCGCGTACCTGAGACCGCGCGGCTGCGC 146811
Qy 247 GCGCGCGCGCGCGGCGCTGCGATGCGACCGCGCGCAGTGTGGGCTGACCATCTCC 306
Db 146810 GCGCGCGCGCGCGGCGCGCTGCGATGCGACCGCGCGCAGTGTGGGCTGACCATCTCC 146751
Qy 307 GCCTACCAAGGAGGACCGCGCTGCTGCGCGCAGTGTGGGCTGCGCGCGCGCTGCTG 366
Db 146750 GCCTACCAAGGAGGACCGCGCTGCTGCGCGCAGTGTGGGCTGCGCGCGCGCTGCTG 146691
Qy 367 TACCGCGCGCGCGCGCTGCGCTCTCATGTGTGTGATGCGCAACCGCGCGCGAGACCTC 426
Db 146690 TACCGCGCGCGCGCGCTGCGCTCTCATGTGTGTGATGCGCAACCGCGCGCGAGACCTC 146631
Qy 427 TACATGTGTCATGTTTCGCGAGGCTTTCCTGACGAGGACCGCGCGCACTGATGTTGG 486
Db 146630 TACATGTGTCATGTTTCGCGAGGCTTTCCTGACGAGGACCGCGCGCACTGATGTTGG 146571
Qy 487 GACCGCAACTACCAACGAGCCTTGGGACCGCGCGCGCGCGGCTGCGCGCGCGGAGCC 546
Db 146570 GACCGCAACTACCAACGAGCCTTGGGACCGCGCGCGCGCGGCTGCGCGCGCGGAGCC 146511
Qy 547 TATCGGAGGTGGAGGCGGAGGATCTTGGGCGGCTGCGAGTGGAGGCGCTGGTCAAGACT 606
Db 146510 TATCGGAGGTGGAGGCGGAGGATCTTGGGCGGCTGCGAGTGGAGGCGCTGGTCAAGACT 146451
Qy 607 GCGAGGTGCTGTGCTGCGCGCAGCGCTGCGGCGCGCAAGCGCGAGTGTATGACAGCC 666
Db 146450 GCGAGGTGCTGTGCTGCGCGCAGCGCTGCGGCGCGCAAGCGCGAGTGTATGACAGCC 146391
Qy 667 TTCAGGCGCTCGAGATTTCGCTGAGTACGTGACAGTCTCT 708
Db 146390 TTCAGGCGCTCGAGATTTCGCTGAGTACGTGACAGTGTAGT 146349

RESULT 8
LOCUS AB097568 2095 bp mRNA linear ROD 28-APR-2004
DEFINITION Rattus norvegicus HAS1 mRNA for hyaluronan synthase 1, complete cds.
ACCESSION AB097568
VERSION AB097568.1 GI:26453348
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
REFERENCE 1
AUTHORS Itano,N., Sawai,T., Ateumi,F., Miyaishi,O., Taniguchi,S., Kannagi,R., Hamaguchi,M. and Kimata,K.
TITLE Selective expression and functional characteristics of three Mammalian hyaluronan synthases in oncogenic malignant transformation
J. Biol. Chem. 279 (18), 18679-18687 (2004)
PUBMED 14724275
REFERENCE 2 (bases 1 to 2095)
AUTHORS Itano,N.
TITLE Direct Submission
Submitted (05-DEC-2002) Naoki Itano, Aichi Medical University,
JOURNAL Institute for Molecular Science of Medicine; Nagakute, Aichi, Aichi

480-1195, Japan (E-mail: itano@amugw.aichi-med-u.ac.jp,
Tel:81-52-264-4811 (ex.2095), Fax:81-561-63-3532)
FEATURES
source location/Qualifiers
1..2095
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
1..2095
/gene="HAS1"
43..1794
/gene="HAS1"
/codon_start=1
/product="hyaluronan synthase 1"
/protein_id="BAC43730.1"
/db_xref="GI:26453349"
/translation="MRQMPKPEAARCCSLARRVLITIFALLILGLMTWAYAAGVP
LADPVGLLAFGLYGLFSLARLLYPRTRVLVMDVGNRAEDLYMDRFEVFADED
LTSAYQEDPTLYRQCLTSARLLYPRTRVLVMDVGNRAEDLYMDRFEVFADED
PATVMDGNTHQWPEAEAGAVEGAYREVEDPGLRAVEALVRRRCVCVAQRWG
KREVMYATFALGDSVDYQVCDSTRLDPMALLELVRLVDLDPKVGNGVDYRLN
PLDSWVSLRLRYWAFNVERACQSYFHCVCISGGLGRLNNLLQGLFLEWNTQKF
LGHTCTGDDRHLLTNRLSMGYATKFTSRSCYSETSSFLRWLSQQTRSKSYFREK
LYNALMHRHAWMTYEAUVGLFPFFVAATVLRFLFYAGRPWALLVLLCVQYALAK
APFAALRGCLRMVLLSLYAPLYMCGLLPAKFLALVTMOSGMCSTGRKLAANYVP
LPALMALLLGLIRSAQEVRAADWSGPRAAEAYHLAAGASAYVAYVWIMLTIIYV
GVRLCRRRSGGRVQV"

ORIGIN

Query Match 63.5%; Score 679.8; DB 10; Length 2095;
Best Local Similarity 84.6%; Pred. No. 6.4e-81;
Matches 794; Conservative 0; Mismatches 127; Indels 18; Gaps 2;
Qy 10 CAGAGCGCGCCAAAGCCACTCTCTGACGCCCGCGCTGCTCCGCGCTTGGCCCGGAGGGTG 69
Db 49 CAGACATGCGCAAGCCCTCAGAGGCGAGCAGTGTGCTGCTCCGCTCTGGCCAGCGGGTG 108
Qy 70 CTGACCATCGCTTCGCCCTGCTCATCTCGGCTCATGACCTGGGCTTACGCCCGCGGG 129
Db 109 CTACGATACCTTCGCCCTGCTCATCTCGGCTCATGACCTGGGCTTACGCCAGCGGA 168
Qy 130 GTGCGCTGCGCTCCGATCGCTACGGCTCTCTGCGCTTTCGCGCTTACGGGCTTCTCTT 189
Db 169 GTACCTCTGCTTCTGATCCTATGGCTCTCTGCGCTTTCGCGCTTATGAGGCTTCTCTC 228
Qy 190 TCAGCGACCTGTGGCGGAGAGCGCTCTTCTGCTACCTGAGGACCGCGGGTGGCGCG 249
Db 229 AGTGGCAGCTAGTGGCAGAGCGCTCTTCTGCTTACCTGAGGACCGAAGGCTGACCGTG 288
Qy 250 GCGGCGGCG-----GGGCGCTGGATGTCAGCCACCGCGCGCAGTGTGGCG 294
Db 289 GCTGCGGCGCGCGCTTTTTCGAAAGGAGCCCTTGGATCGGCGCACTGCGCGCAGCGTGGCA 348
Qy 295 CTGACCATCTCCGCTTACCAAGAGGACCCCGCGTACCTTGGCGCCAGTGCCTGGCGTCCGCC 354
Db 349 CTACCATCTCAGCTTACCAAGAGGACCCCACTTACCTGCGCGCAGTGTGACCTCCGCG 408
Qy 355 CGCGCTCTGTGTACCGCGCGCGCGCTGCGCTCTCTATGTGTGTGTGTGTGTGTGTGTGTGT 414
Db 409 CGCGCTCTGTGTACCGCGCGCGCGCTGCGCTCTCTATGTGTGTGTGTGTGTGTGTGTGTGT 468
Qy 415 GCGGAGGACCTCTACATGTCGATGTCGCGAGGCTTTCGCTGACGAGGACCCCGCC 474
Db 469 GCGGAGGATCTGTACATGTCGAGATGTCGAGAGTCTTCGAGAGTATGAGACCTTGCC 528
Qy 475 ACGTACGTGTGGAGCGGCAACTACCACTGAGCCCTTGGGAACCCCGC---GGCGCGCGGCGCG 531
Db 529 ACTTATGTGTGGATGGCAACTACCATCAGCTTTGGGAGCCAGCTGAGGCGCGGGTGT 588
Qy 532 GTGGCGCGCGGAGCGCTTATCGGAGGTGGAGCGGAGGATCTTGGCGCGCTGGCAGTGGAG 591
Db 589 GTGGGTGAAGTGTGCTACCGCGAGGTGGAGGCTGAGGACCTTGGCGCGCTGGCGGTAGAG 648
Qy 592 GCGCTGTGTAGGACTCGCAGGTGCTGTGCTGCGCGCAGCGCTTGGGCGCGCAAGCGCGAG 651

[illegible]

RESULT 9	El3681	2102 bp	DNA	linear	PAT 27-APR-1999
LOCUS	El3681				
DEFINITION	DNA encoding novel mouse hyaluronate synthetase.				
ACCESSION	El3681				
VERSION	El3681.1	GI:3252450			
KEYWORDS	JP 1997224674-A/1.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. 1 (bases 1 to 2102)				
AUTHORS	Itano, N. and Kimata, H.				
TITLE	POLYPEPTIDE OF NEW HYALURONIC ACID-SYNTHETIC ENZYME AND DNA CODING THE SAME				
JOURNAL	Patent: JP 1997224674-A 1 02-SEP-1997;				

```

ORIGIN
/adm_xref="taxon:U0090"
Query Match      63.3%   Score 678.2;   DB 6;   Length 2102;
Best Local Similarity 84.5%;   Pred. No. 1e-80;
Matches 793;   Conservative 0;   Mismatches 128;   Indels 18;   Gaps 2;

```

ORGANISM	Mus sp.	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	Yoichi, Y., Naoki, I. and Koji, K.	
TITLE	Hyaluronate synthase promoter DNA	
JOURNAL	Patent: JP 1999196875-A 2 27-JUL-1999;	
COMMENT	SEIKAGAKU KOGYO CO LTD	
OS	Mus sp. (mouse)	
PN	JP 1999196875-A/2	
PD	27-JUL-1999	
PF	14-JAN-1998 JP 1998006191	
PR		
PI	YOICHI YAMADA, NAOKI ITANO, KOJI KIMATA	
PC	C12N15/09, C12N9/00, C12Q1/68// (C12N15/09, C12R1:91), C12N15/00,	
PD	(C12N15/00, C12R1:91)	
CC	Strandedness: Double;	
CC	Topology: Linear;	
FH	Key	
FT	CDS	
FEATURES	Location/Qualifiers	
source	1..2102	
	/organism="Mus sp."	
	/mol_type="genomic DNA"	
	/db_xref="taxon:10095"	
ORIGIN		
Query Match	63.3%; Score 678.2; DB 6; Length 2102;	
Best Local Similarity	84.5%; Pred. No. 1e-80;	
Matches	793; Conservative 0; Mismatches 128; Indels 18; Gaps 2;	
Qy	10 CAGACGGCCCAAGCCACTCTGCAGCCGCGCTGCTCGGCTGGCCGCGAGGGTG 69	
Db	55 CAGACATGCGCAAGCCCTCAGAGCAGCGGTGCTGCTGCGCTGGCCGCGAGCA 114	
Qy	70 CTGACCATCGCTTCCGCTCTCATCTCGGCTCATGACTGCGCTACGCGCGGG 129	
Db	115 CTCAGATCATCTTTGCCCTGCTCATCTGGGCTCATGACCTGGGCTACGCCGAGGC 174	
Qy	130 GTGCGCTGGCTCCGATCGCTACGGCTCTCGGCTTGGCTTACGGGGCTTCTT 189	
Db	175 GTTCTCTGGCTTCAGATCGCTATGGAATCTCTGGCTTTGGCTTATGGGGCATTCCTC 234	
Qy	190 TCAGGGACCTGTGGCGCAGAGCTCTTCGGTACCTGGAGCAGCGCGGGTGGCGGC 249	
Db	235 AGCGCACACCTAGTGGCAGAGCTCTTCCTTACCTGGAGCAGCGAGGGTGGCAGCG 294	
Qy	250 GCGGCGCG-----GGGGCGCTGGATGCGACGACCGCGCGCAGTGTGGCG 294	
Db	295 GCTGGCGCGCTCCTTGGCGAAGGGCCCTGGATGGGCGACTGACGAGCGGTGGCA 354	
Qy	295 CTGACCATCTCCGCTACACAGAGAGACCCCGCTACCTGCGCCAGTGCCTGGCGTCCGCC 354	
Db	355 CTCACCATCTCAGCTACCAAGAGAGTCCCGCTTACCTGCGCCAGTGTTCGACCTCGCG 414	
Qy	355 CGCGCCCTGCTGTACCGCGCGCGGCTCGGCTCTCTCATGTTGGTGGATGGCAACGCG 414	
Db	415 CGCGCCCTGCTGTACCGCGCACACAGAGTTACGCGTGTCTCATGGTGGAGCGCAACGCG 474	
Qy	415 GCGGAGGACCTCTACATGGTGCAGATGTTCCGCGAGGTCTTCGCTGACGAGAGCCCGCGC 474	
Db	475 GCTGAGGATCTGTATATGTTGACATGTTCCGAGAGTCTTCGCGATGAGGACCCCGCC 534	
Qy	475 ACGTACGTGTGGGACGGCAATACCAACAGCCCTGGGAACCCCGCGCG-----CGGGCGCG 531	
Db	535 ACTTATGTGGGATGGCAATCAACATCAGCCCTGGGAACCCAGCGAGGCTACCGGGCT 594	
Qy	532 GTGGCGCGGAGCCTATCGGGAGGTGAGCGGAGGATCTTGGCGGCTGGCAGTGGAG 591	
Db	595 GTCGGTGAAGTGGCTTACCGGAGGTGGAGCGAGGACCCCGCGGTTGGCGGTGGAG 654	
Qy	592 GCGCTGTGTAGGACTCGCAGGTGTGTGCTGTGCGCAGCGCTCGGGCGGCAAGCGGAG 651	
Db		

Db	655 GCGCTGTGTAGAACACACGCGAGGTGCGTGTGCTGCGCTCAGCGCTTGGGGCGCAACGTCAG 714	
Qy	652 GTCATGTACACAGCCTTCAAGCGCTCGGAGATTGGTGGACTAGTGCAGGCTCTGTGAC 711	
Db	715 GTCATGTACACAGCTTTCAAGGCACTCGGGCACTCCGTGGACTACGTGCAGGCTCTGTGAC 774	
Qy	712 TCGGACACAAGGTTGGACCCCATCGCATCTGCTGAGCTCGTGGGCTACTGGACGAGGAC 771	
Db	775 TCAGACACAAGACTAGACCCCATGCACTGCTGAGCTGTGCGAGTGTGGATGGAAGAC 834	
Qy	772 CCCCGGTAGGGCTGTGTGGGAGCTGCGGATCCCTTAAACCTCTGGACTCTCTGGGTC 831	
Db	835 CCCCGGTAGGGCTGTGTGGAGGGATGTAGGATCCTTAAACCTCTGGACTCTCTGGGTC 894	
Qy	832 AGCTTCTTAGCAGCCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGCTTGTGACAGG 891	
Db	895 AGCTTCTTAGCAGCTCTTCGATACTGGGTAGCCTTCAATGTGGAAACGAGCTTGTGAGAGC 954	
Qy	892 TACTTCCACTGTGTATCTGTCATCAGCGGTTCTCTAGST 930	
Db	955 TACTTCCACTGTGTGCTCTGTCATCAGTGTCTCTGGST 993	
RESULT	11	
E30971		
LOCUS	E30971 2102 bp DNA linear PAT 18-JUN-2001	
DEFINITION	Hyaluronic acid synthetase modified protein.	
ACCESSION	E30971	
VERSION	E30971.1 GI:13017286	
KEYWORDS	JP 2000004886-A/1.	
SOURCE	Mus sp.	
ORGANISM	Mus sp.	
REFERENCE	1 (bases 1 to 2102)	
AUTHORS	Naoki, I., Mamoru, Y. and Koji, K.	
TITLE	Hyaluronic acid synthetase modified protein	
JOURNAL	Patent: JP 2000004886-A 1 11-JAN-2000;	
COMMENT	SEIKAGAKU KOGYO CO LTD	
OS	Mus sp. (mouse)	
PN	JP 2000004886-A/1	
PD	11-JAN-2000	
PF	24-JUN-1998 JP 1998193788	
PR		
PI	NAOKI ITANO, MAMORU YOSHIDA, KOJI KIMATA	
PC	C12N15/09, C12N9/00, C12N15/00	
CC	Strandedness: Double;	
CC	Topology: Linear;	
FH	Key	
FT	CDS	
FEATURES	Location/Qualifiers	
source	1..2102	
	/organism="Mus sp."	
	/mol_type="genomic DNA"	
	/db_xref="taxon:10095"	
ORIGIN		
Query Match	63.3%; Score 678.2; DB 6; Length 2102;	
Best Local Similarity	84.5%; Pred. No. 1e-80;	
Matches	793; Conservative 0; Mismatches 128; Indels 18; Gaps 2;	
Qy	10 CAGACGGCCCAAGCCACTCTGCAGCCGCGCTGCTCGGCTGGCCGCGAGGGTG 69	
Db	55 CAGACATGCGCAAGCCCTCAGAGCAGCGGTGCTGCTGCGCTGGCCGCGAGCA 114	
Qy	70 CTGACCATCGCTTCCGCTCTCATCTCGGCTCATGACTGCGCTACGCGCGGG 129	
Db	115 CTCAGATCATCTTTGCCCTGCTCATCTGGGCTCATGACTGCGCTACGCCGAGGC 174	
Qy	130 GTGCGCTGGCTCCGATCGCTACGGCTCTCGGCTTGGCTTACGGGGCTTCTT 189	
Db	175 GTTCTCTGGCTTCAGATCGCTATGGAATCTCTGGCTTTGGCTTATGGGGCATTCCTC 234	
Qy	190 TCAGGGACCTGTGGCGCAGAGCTCTTCGGTACCTGGAGCAGCGCGGGTGGCGGC 249	
Db	235 AGCGCACACCTAGTGGCAGAGCTCTTCCTTACCTGGAGCAGCGAGGGTGGCAGCG 294	
Qy	250 GCGGCGCG-----GGGGCGCTGGATGCGACGACCGCGCGCAGTGTGGCG 294	
Db	295 GCTGGCGCGCTCCTTGGCGAAGGGCCCTGGATGGGCGACTGACGAGCGGTGGCA 354	
Qy	295 CTGACCATCTCCGCTACACAGAGAGACCCCGCTACCTGCGCCAGTGCCTGGCGTCCGCC 354	
Db	355 CTCACCATCTCAGCTACCAAGAGAGTCCCGCTTACCTGCGCCAGTGTTCGACCTCGCG 414	
Qy	355 CGCGCCCTGCTGTACCGCGCGCGGCTCGGCTCTCTCATGTTGGTGGATGGCAACGCG 414	
Db	415 CGCGCCCTGCTGTACCGCGCACACAGAGTTACGCGTGTCTCATGGTGGAGCGCAACGCG 474	
Qy	415 GCGGAGGACCTCTACATGGTGCAGATGTTCCGCGAGGTCTTCGCTGACGAGAGCCCGCGC 474	
Db	475 GCTGAGGATCTGTATATGTTGACATGTTCCGAGAGTCTTCGCGATGAGGACCCCGCC 534	
Qy	475 ACGTACGTGTGGGACGGCAATACCAACAGCCCTGGGAACCCCGCGCG-----CGGGCGCG 531	
Db	535 ACTTATGTGGGATGGCAATCAACATCAGCCCTGGGAACCCAGCGAGGCTACCGGGCT 594	
Qy	532 GTGGCGCGGAGCCTATCGGGAGGTGAGCGGAGGATCTTGGCGGCTGGCAGTGGAG 591	
Db	595 GTCGGTGAAGTGGCTTACCGGAGGTGGAGCGAGGACCCCGCGGTTGGCGGTGGAG 654	
Qy	592 GCGCTGTGTAGGACTCGCAGGTGTGTGCTGTGCGCAGCGCTCGGGCGGCAAGCGGAG 651	
Db		

QY 190 TCAGCGCACCTGCTGGCGCAGAGCCTCTTCCGCTACCTGGAGCAGCGCGCGGTGGCGGG 249
Db 235 AGCGCACACCTAGTGGCAGAGCCTCTTCCGCTTACCTGGAGCAGCGAGGGTGGCAGCG 294
QY 250 GCGGCGCG-----GGGGCGCTGGATGACAGCCACCGCGCGCAGTGTGGCG 294
Db 295 GCTGGCGGCGCTCTTGGCGAAGGGGCCCTGGATGGCCACCTGACGCGAGCGTGGCA 354
QY 295 CTGACCATCTCCGCTTACAGAGAGACCCCGGCTACCTGCGCCAGTGTGGCGTCCGCC 354
Db 355 CTCACCATCTCAGCCTTACCAAGAGATCCCGCTTACCTGCGCCAGTGTGACCTCCGCG 414
QY 355 CGCGCCCTCTGTACCGCGCGCGCGCTGCGCGCTCTCATGGTGGTGGATGGCAACCGC 414
Db 415 CGCGCCCTCTGTACCGCGCACAGAGTTACCGCTGCTCATGGTGGTGGCAACCGC 474
QY 415 GCGGAGGACCTTACATGGTGCACATGTTCCGCGAGGCTCTCGCTGACGAGACCCCGC 474
Db 475 GCTGAGGATCTGTATCGATGGATGCGCATGTTCCGAGAGTCTTCCGATGAGGACCCCGC 534
QY 475 ACCTGATGTGGATGGCAATACCATCAGCCCTGGGAACCAAGCGCTACGCGCGCT 531
Db 535 ACTTATGTGGATGGCAATACCATCAGCCCTGGGAACCAAGCGCTACGCGCGCT 594
QY 532 GTGGCGCGCGAGCCTATCGGAGGTGAGGCGGAGGATCTTGGCGCGCTGGCAGTGGAG 591
Db 595 GTCGGTGAAGGTGCTTACCGGAGGTGGAGCGGAGGACCCCGCGGCTTGGCGGTGGAG 654
QY 592 GCGCTGTGAGGACTCGCAGGTGCTGTGCGGCGCAGCGCTGGGCGGCGCAAGCGCGAG 651
Db 655 GCGCTGTGAGGACTCGCAGGTGCTGTGCGGCGCAGCGCTGGGCGGCGCAAGCGTGGAG 714
QY 652 GTCTGTGTACACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTAGTGCAGGTCTGTGAC 711
Db 715 GTCTGTGTACACAGCTTCAAGGCACTGGGCGACTCGGTGGACTAGTGCAGGTCTGTGAC 774
QY 712 TCGGACACAGGTGGACCCCATGCACTGCTGGAGCTCGTGGCGGCTTGTGAGCAGGAC 771
Db 775 TCAGACACAGACTAGACCCCATGCACTGCTGGAGCTTGTGGCGGCGCAAGCGTGGAG 834
QY 772 CCGCGGTAGGGGCTGTGGTGGGACGTCGGGATCTTAACTTGGAGCTTGTGGAGTGGAG 831
Db 835 CCGCGGTAGGGGCTGTGGAGGGGATGTGAGGATCTTAACTTGGAGCTTGTGGAGTGGAG 894
QY 832 AGCTTCTTAAGCAGCTCGGATPACTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGC 891
Db 895 AGCTTCTTAAGCAGCTTTCGATCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGC 954
QY 892 TACTTCCATGTGTATCTGCAATCAGCGGTTCTTAGGT 930
Db 955 TACTTCCATGTGTCTGTCATCAGTGTCTCTGGGT 993

RESULT 12
E34326
LOCUS 2102 bp DNA linear PAT 31-JAN-2002
DEFINITION DNA for gene targeting hyaluronic acid synthase gene.
ACCESSION E34326
VERSION E34326.1 GI:18624311
KEYWORDS JP 2000116382-A/1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2102)
AUTHORS Itano, N. and Kimata, K.
TITLE DNA for gene targeting hyaluronic acid synthase gene
JOURNAL Patent: JP 2000116382-A 1 25-APR-2000;
SEIKAGAKU KOGYO CO LTD
COMMENT OS Mus musculus (mouse)
PN JP 2000116382-A/1
PD 25-APR-2000
PF 13-OCT-1998 JP 1998291201

PR NAOKI ITANO, KOJI KIMATA
PI C12N15/09, C12N5/10, G01N33/50//C12N9/00, C12Q1/68, (C12N15/09, PC
PC C12R1:91),
PC (C12N5/10, C12R1:91), C12N15/00, C12N5/00, (C12N15/00, C12R1:91),
PC (C12N5/00, C12R1:91)
CC
FH Key Location/Qualifiers
FT CDS Location/Qualifiers
1..2102
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
FEATURES
source
ORIGIN
Query Match 63.3%; Score 678.2; DB 6; Length 2102;
Best Local Similarity 84.5%; Pred. No. 1e-80;
Matches 793; Conservative 0; Mismatches 128; Indels 18; Gaps 2;
QY 10 CAGGACGCGCCCAAGCCCACTCTCTGACGCCCGCGCTCTCGGCGCTGGCCCGGAGGTG 69
Db 55 CAGGACATGCCAAAGCCCTCAGAGCAGCGGCTTGTCTCTGGCTTGGCCAGCGAGCA 114
QY 70 CTGACCATCGGCTTCGCCCTGCTCATCTGGGCTCATGACTGGGCTTACGCGCGCGG 129
Db 115 CTCAGATCATCTTTGGCTCTCATCTGGGCTTGTGCTCTTGGGCTTGTGCGCATTCCTC 174
QY 130 GTGCGCTGGGCTCGATCGCTACGGCTCTCTGGCTTCTGGCTTACGCGGCTTCTCTT 189
Db 175 GTTCTCTGGGCTCAGATCGCTATGGACTCTCTGGCTTGTGCTCTTGTGGGCTTCTCTC 234
QY 190 TCAGCGCACCTGCTGGCGCAGAGCCTCTTCCGCTACCTGGAGCAGCGCGGTGGCGGG 249
Db 235 AGCGCACACCTAGTGGCAGAGCCTCTTCCGCTTACCTGGAGCAGCGAGGTGGCAGCG 294
QY 250 GCGGCGCG-----GGGGCGCTGGATGACAGCCACCGCGCGCAGTGTGGCG 294
Db 295 GCTGCGCGCGCTCTCTTGGCGAAGGGGCCCTTGGATGGCGGCACTTGCACGACGCTGGCA 354
QY 295 CTGACCATCTCGGCTTACAGAGAGACCCCGGCTACCTGCGCGCTGGAGTGGCTGGCGCG 354
Db 355 CTCACCATCTCAGCCTTACAGAGAGATCCCGCTTACCTGCGCGAGTGTGTGACCTCCGCG 414
QY 355 CGCGCCCTCTGTATACCGCGCGCGCTGGCGCTCTCATGGTGGTGGATGGCAACCGC 414
Db 415 CGCGCTTGTGTATACCGCGCACAGAGTTACGCTGCTCATGGTGGTGGAGCGGCAACCGC 474
QY 415 GCCAGGACCTCTACATGGTTCGATGTTCCGCGAGGTTCTTCTGCTGACGAGACCCCGC 474
Db 475 GCTGAGGATCTGTATCGTGGACATGTTCCGAGAGTCTTCCGCGATGAGAGCCCGCGC 534
QY 475 ACGTACGTGTGGAGCGGCACTTACACAGCCCTGGGNAACCCCGCGCGG---CGGCGCGG 531
Db 535 ACTTATGTGTGGATGGCAACTTACATCAGCCCTTGGGAACCAAGCGGCTTACGCGCGCT 594
QY 532 GTGGCGCGCGAGCCTATCGGAGAGTGGAGCGGAGGATCTTGGCGCGCTTGGCAGTGGAG 591
Db 595 GTCGGTGAAGGTGCTTACCGGAGAGTGGAGCGGAGGACCCCGCGGCTTGGCGGTGGAG 654
QY 592 CGCTGTGTGAGGACTCGCAGGTGCTGTGCTGGCGCAGCGCTGGGCGGCGCAAGCGCGAG 651
Db 655 GCGCTGTGTGAGAACACGCGAGGTGCTGTGCTGGCTCAGCGTTGGGCGGCGCAACGCTGAG 714
QY 652 GTCATGTACACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTAGTGCAGGTCTGTGAC 711
Db 715 GTCATGTACACAGCTTTCAGGSCACTGGGCGACTCCCGTGGACTAGTGCAGGTCTGTGAC 774
QY 712 TCGGACACAGGTGGACCCCATGCACTGCTGGAGCTCGTGGCGGCTTGTGAGCAGGAC 771
Db 775 TCAGACACAGACTAGACCCCATGCACTGCTGGAGCTTGTGGAGTGTGGATGAGAGAC 834
QY 772 CCGCGGTAGGGGCTGTGGTGGGACGTCGGGATCTTAACTTGGAGCTTGTGGAGTGGAG 831

AUTHORS

Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamen, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Cariaga, K., Coleman, B., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laric, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Mastello, C., Maskeri, B., McDowell, J., Padurigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schandler, K., Schueler, M.G., Sison, C., Stantripop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.

TITLE

NISC Comparative Sequencing Initiative

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 193986)

AUTHORS

Green, E.D.

TITLE

Direct Submission

JOURNAL

Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717

REFERENCE

3 (bases 1 to 193986)

AUTHORS

Green, E.D.

TITLE

Direct Submission

JOURNAL

Submitted (19-NOV-2002) NIH Intramural Sequencing Center, 8717

COMMENT

Submitted (19-NOV-2002) NIH Intramural Sequencing Center, 8717

Governmont Circle, Gaithersburg, MD 20877, USA

On Nov 19, 2002 this sequence version replaced gi:22218453.

----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc_zoo@nhgri.nih.gov

----- Project Information

Center project name: dxd

Center clone name: 426A12

The sequence data in this record represents an 'enhanced' version of a phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 191764 bases at least Q40

Consensus quality: 192510 bases at least Q30

Consensus quality: 192998 bases at least Q20

Insert size: 184000; agarose-fp

Insert size: 193286; sum-of-contigs

Quality coverage: 11.49x in Q20 bases; agarose-fp

Quality coverage: 10.94x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 11512: contig of 11512 bp in length

* 11513 11612: gap of unknown length

* 11613 59488: contig of 47876 bp in length

* 59489 59588: gap of unknown length

* 59589 103266: contig of 43678 bp in length

* 103267 103366: gap of unknown length

* 103367 127885: contig of 24519 bp in length

* 127886 127985: gap of unknown length

* 127986

* 129692: contig of 1707 bp in length

* 129693 129792: gap of unknown length

* 129793 131077: contig of 1285 bp in length

* 131078 131177: gap of unknown length

* 131178 174663: contig of 43486 bp in length

* 174664 174763: gap of unknown length

* 174764 193986: contig of 19223 bp in length.

Location/Qualifiers

source

1. 193986

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

/clone="CH251-426A12"

/clone_lib="CH251"

1. 11512

/note="assembly_fragment"

clone_end:SP6

vector_side:left"

1. 176

/note="clone overlaps with GenBank Accession Number

AC132069 clone CH251-132A6 (center project name dxc)"

11613. .59488

/note="assembly_fragment"

59589. .103266

/note="assembly_fragment"

103367. .127885

/note="assembly_fragment"

127986. .129692

/note="assembly_fragment"

129793. .131077

/note="assembly_fragment"

131178. .174663

/note="assembly_fragment"

174764. .193986

/note="assembly_fragment"

clone_end:T7

vector_side:right"

ORIGIN

```

Query Match      62.5%; Score 669; DB 2; Length 193986;
Best Local Similarity 97.2%; Pred. No. 4.9e-80;
Matches 697; Conservative 0; Mismatches 5; Indels 15; Gaps 1;

Qy  7  CAGCAGGACGGCGCCCAAGCCCACTCTGCAGCCCGCCGCTCTCGCGGCTGCGCCGAGG 66
Db  132828  CAGCAGGACGGCGCCCAAGCCCACTCTGCAGCCCGCCGCTCTCGCGGCTGCGCCGAGG 132769

Qy  67  GTGCTGACCATCGGCTTTCGCGCTGCTATCTCTGGGCTTCATGACCTGGGCTTACGCGCC 126
Db  132768  GTGCTGACCATCGGCTTTCGCGCTGCTATCTCTGGGCTTCATGACCTGGGCTTACGCGCC 132709

Qy  127  GGGGTGCGGCTGGCTCCGATCGGTACGGCTCTCGGCTTTCGGCTTTCAGGGGCTTC 186
Db  132708  GGGGTGCGGCTGGCTCCGATCGGTACGGCTTTCGGCTTTCAGGGGCTTC 132649

Qy  187  CTTTCAGCGCACCTGGTGGCGCAGAGCCTCTTCGGCTACCTGGAGCACCGCGGCTGGCG 246
Db  132648  CTTTCAGCGCACCTGGTGGCGCAGAGCCTCTTCGGCTACCTGGAGCACCGCGGCTGGCG 132589

Qy  247  -----GCGGCGCGCGGGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGG 291
Db  132588  GCGGCGCGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGCGGCGGCGGCGG 132529

Qy  292  GCGGTGACCATCTCGGCTTACAGAGGAGACCGCGGTACCTCGCGGCTGCTGCGGCTTC 351
Db  132528  GCGGTGACCATCTCGGCTTACAGAGGAGACCGCGGTACCTCGCGGCTGCTGCGGCTTC 132469

Qy  352  GCCCGCGGCTGCTGTACCCCGCGCGCGGCTGCGGCTTCCTCATGGTGGTGGTGGTGG 411
Db  132468  GCCCGCGGCTGCTGTGTACCCCGCGCGCGGCTGCGGCTTCCTCATGGTGGTGGTGGTGG 132409

Qy  412  CGCGCGGAGGACCTCTTACATCGTGCACATGTTTCGCGGAGGCTTTCGCTGACGAGACCCC 471

```

Db 132408 CGCGCCGAGGACCTCTCATGCTGCACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCC 132349

QY 472 GCCAGTACGTGTGGGACGGCAACTACACACAGCCCTCGGAAACCCGCGCGCGCGCGCGCG 531

Db 132348 GCCAGTACGTGTGGGACGGCAACTACACACAGCCCTCGGAAACCCGCGCGCGCGCGCGCG 132289

QY 532 GTGGCGCGCGGAGCCTATCTCGGAGGTGAGCGGAGGATCCTGGGCGCTGCGCAGTGGAG 591

Db 132288 GTGGCGCGCGGAGCCTACCGGAGGTGAGCGGAGGATCCTGGGCGCTGCGCAGTGGAG 132229

QY 592 GCGCTGCTGAGGACTCGCAGGTGCGTGTGCTGGCGCAGCGCTGGGCGGCAAGCGCAG 651

Db 132228 GCGCTGCTGAGGACTCGCAGGTGCGTGTGCTGGCGCAGCGCTGGGCGGCAAGCGCAG 132169

QY 652 GTCATGTACACAGCCTTCAAGCGCTCGGAGATTGCGTGGACTAGCTGCAGGTCTGT 708

Db 132168 GTCATGTACACAGCCTTCAAGCGCTCGGAGATTGCGTGGACTAGCTGCAGGTCTGT 132112

RESULT 15 AC137058/c

LOCUS AC137058.3 185623 bp DNA linear HTG 21-FEB-2003

DEFINITION Papio anubis clone RP41-126M5, WORKING DRAFT SEQUENCE, 10 ordered

ACCESSION AC137058

VERSION AC137058.3 GI:28460766

KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.

SOURCE Papio anubis (olive baboon)

ORGANISM Papio anubis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopithecinae; Papio.

REFERENCE AUTHORS

1. (bases 1 to 185623)
Akter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Hachichi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Latic,P., Lee-Lin,S.-Q., Legaapi,R., Maduro,Q.L., Maduro,V.B.,
Margulies,E.H., Mastello,C., Maskeri,B., McDownell,J.,
Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Weherby,K.D., Wiggins,L., Young,A. and Green,E.D.

NISC Comparative Sequencing Initiative

TITLE JOURNAL

REFERENCE 2 (bases 1 to 185623)

AUTHORS Green,E.D.

TITLE Direct Submission

JOURNAL Submitted (15-NOV-2002) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA

REFERENCE AUTHORS

TITLE Direct Submission

JOURNAL Submitted (21-FEB-2003) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA

On Feb 21, 2003 this sequence version replaced gi:27476124.

COMMENT

----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoonhgri.nih.gov
----- Project Information
Center project name: dtj
Center clone name: 126M05

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average

coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 183092 bases at least Q40
Consensus quality: 183851 bases at least Q30
Consensus quality: 184434 bases at least Q20
Insert size: 160000; agarose-fp
Quality coverage: 14.13x in Q20 bases; agarose-fp
Quality coverage: 12.24x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and the accession number will be preserved.

* 1 9984: contig of 9984 bp in length

* 9985 10084: gap of unknown length

* 10085 49805: contig of 39721 bp in length

* 49806 49905: gap of unknown length

* 49906 66936: contig of 17031 bp in length

* 66937 67036: gap of unknown length

* 67037 83604: contig of 16568 bp in length

* 83605 83704: gap of unknown length

* 83705 108246: contig of 24542 bp in length

* 108247 108346: gap of unknown length

* 108347 112021: contig of 3675 bp in length

* 112022 112121: gap of unknown length

* 112122 134817: contig of 22696 bp in length

* 134818 134917: gap of unknown length

* 134918 147290: contig of 12373 bp in length

* 147291 147390: gap of unknown length

* 147391 185337: contig of 37947 bp in length

* 185338 185437: gap of unknown length

* 185438 185623: contig of 186 bp in length.

FEATURES source

1. 185623
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
/clone="RP41-126M5"
/clone_lbb="RP41"

1. 15263

/note="clone overlaps with GenBank Accession Number AC130273 clone RP41-30706 (center project name dsu)"

1. 9984

/note="assembly_fragment
clone_end:177
vector_side:left"

10085. 49805

/note="assembly_fragment"

49906. 66936

/note="assembly_fragment"

67037. 83604

/note="assembly_fragment"

83705. 108246

/note="assembly_fragment"

108347. 112021

/note="assembly_fragment"

112122. 134817

/note="assembly_fragment"

134918. 147290

/note="assembly_fragment"

147391. 185337

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

misc_feature

/note="assembly_fragment"
185438..185623
/note="assembly_fragment
clone_end:SP6
vector_side:right"

misc_feature

ORIGIN

```
Query Match      59.3%; Score 635.4; DB 2; Length 185623;
Beat Local Similarity 94.3%; Pred. No. 1.4e-75;
Matches 676; Conservative 0; Mismatches 26; Indels 15; Gaps 1;

Qy 7 CAGCAGGAGCGCCCAAGCCACTCTGCAGCCCGCTGCTCGGCTGCGCCGCGAGG 66
Db 138595 CAGCGGACACCCCAAGCCACTCTGCAGCCCGCTGCTCGGCTGCGCCGCGAGG 138536

Qy 67 GTGCTGACCATCGCCTTGCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCCGCC 126
Db 138535 GTGCTGACCATCGCCTTGCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCCGCC 138476

Qy 127 GGGGTGCGCTGGCTTCCGATCGCTACCGGCTCTCTGGGCTTCTGGGCTTCTACCGGGCCTTC 186
Db 138475 GGGGTGCGCTGGCTTCCGATCGCTACCGGCTCTCTGGGCTTCTGGGCTTCTACCGGGCCTTC 138416

Qy 187 CTTTCAGCGCACCTGGTGGCGAGAGCCTCTTCGGGTACTGGAGACCGGCGGGTGGG 246
Db 138415 CTTTCAGCGCACCTGGTGGCGAGAGCCTCTTCGGGTACTGGAGACCGGCGGGTGGG 138356

Qy 247 GCGGCGCGG-----CGGGGCGCTGGGATGACGACCGCGCGCAGTGTG 291
Db 138355 GCGGCGCGGCGGCGCGGCGCACGGGGGCGCTGGATGACGACCGCGCGCAGCGTG 138296

Qy 292 GCGCTGACATCTCGCTTACAGAGGACCCCGCTACCTGCGCCAGTGCCTGCGCTCC 351
Db 138295 GCGCTGACATCTCGCTTACAGAGGACCCCGCTACCTGCGCCAGTGCCTGCGCTCC 138236

Qy 352 GCGGCGCGCTGCTGTACCGCGCGCGGCTGCGGCTCTCTCATGGTGGTGGATGGCAAC 411
Db 138235 GCGGCGCGCTGCTGTACCGCGCGCGGCTGCGGCTCTCTCATGGTGGTGGATGGCAAC 138176

Qy 412 CGCGCGGAGGACCTCTACATGGTGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCC 471
Db 138175 CGCGCGGAGGACCTCTACATGGTGACATGTTCCGCGAGGTCTTCGCGAGGACCCC 138116

Qy 472 GCCAGTACGTGTGGGACGGCAACTACCAAGCCCTGGGAACCGCGCGCGCGCGCG 531
Db 138115 GCCAGTACGTGTGGGACGGCAACTACCAAGCCCTGGGAACCGCGCGCGCGCGCG 138056

Qy 532 GTGGCGCGGAGCCTATCGGAGGTGGAGCGGAGGATCTTGGGCGGCTGGCAGTGGAG 591
Db 138055 GTGGCGCGGAGCCTATCGGAGGTGGAGCGGAGGATCTTGGGCGGCTGGCAGTGGAG 137996

Qy 592 GCGCTGGTGAGGACTCGCAGGTGCTGTGCGGCGCAGCGCTGGGCGGCGAAGCGCGAG 651
Db 137995 GCGCTGGTGAGGACTCGCAGGTGCTGTGCGGCGCAGCGCTGGGCGGCGAAGCGCGAG 137936

Qy 652 GTCATGTACAGCGCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGAGGTCTGT 708
Db 137935 GTCATGTACAGCGCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGAGGTCTGT 137879
```

Search completed: March 13, 2005, 08:01:29
Job time : 5065.4 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 02:13:47 ; Search time 5115.08 Seconds
(without alignments)
10259.279 Million cell updates/sec

Title: US-10-672-399-5

Perfect score: 1083

Sequence: 1 atgagacagcaggacgcgc.....tccaagtcgtactccgtga 1083

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb:ba.*

2: gb:htg.*

3: gb:in.*

4: gb:om.*

5: gb:ov.*

6: gb:pat.*

7: gb:ph.*

8: gb:pl.*

9: gb:pr.*

10: gb:ro.*

11: gb:sts.*

12: gb:sv.*

13: gb:un.*

14: gb:vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	933	86.1	2088	9 HSU59269	U59269 Human hyalu
2	933	86.1	2116	6 AR220003	AR220003 Sequence
3	920.8	85.0	2119	9 BC035837	BC035837 Homo sapi
4	920.6	85.0	2108	9 HUMHAS	D84424 Homo sapien
5	920.6	85.0	2117	6 AR137038	AR137038 Sequence
6	855	78.9	2109	9 AY463695	AY463695 Papio anu
7	697.2	64.4	229155	9 AC018755	AC018755 Homo sapi
8	679.4	62.7	2095	10 AB097568	AB097568 Rattus no
9	677.8	62.6	2102	6 E13681	E13681 DNA encodin
10	677.8	62.6	2102	6 E28454	E28454 Hyaluronate
11	677.8	62.6	2102	6 E30971	E30971 Hyaluronon
12	677.8	62.6	2102	6 E34326	E34326 DNA for gen
13	677.8	62.6	2102	10 MUSHAS	D82964 Mus musculu
14	669	61.8	193986	2 AC130783	AC130783 Pan trogl
15	635.4	58.7	185623	2 AC137058	AC137058 Papio anu
16	620.4	57.3	1581	6 CQ731079	CQ731079 Sequence
17	492	45.4	293184	2 AC079487	AC079487 Mus muscu
18	490.8	45.3	72955	2 AC108651	Continuation (4 of
19	490.8	45.3	257325	2 AC116203	AC116203 Rattus no

BV099313 RPAMMSQ0
 BV160501 RPAMMSQ0
 AY437407 Danio rer
 AB055979 Oryctolag
 BV161377 RPAMMSQ0
 CQ723590 Sequence
 CQ723590 Sequence
 AF234839 Homo sapi
 AF232772 Homo sapi
 CQ776471 Sequence
 AB159675 Sus scrof
 BC047963 Xenopus l
 M22249 X.laavis DG
 AB097569 Rattus no
 AF015780 Xenopus l
 BC077983 Xenopus l
 E30973 Hyaluronon
 AR373269 Sequence
 CQ777380 Sequence
 UB6408 Mus musculu
 CQ776472 Sequence
 BC021853 Homo sapi
 BV161384 RPAMMSQ0
 BX539313 Zebrafish
 BV160507 RPAMMSQ0
 BV161386 RPAMMSQ0
 BV099322 RPAMMSQ0

ALIGNMENTS

RESULT 1
 HSU59269
 LOCUS Human hyaluronan synthase mRNA, complete cds.
 DEFINITION
 ACCESSION U59269
 VERSION U59269.1 GI:1556464
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 2088)
 AUTHORS Shyjan,A.M., Heidin,P., Butcher,E.C., Yoshino,T. and Briskin,M.J.
 TITLE Functional cloning of the cDNA for a human hyaluronan synthase
 JOURNAL J. Biol. Chem. 271 (38), 23395-23399 (1996)
 MEDLINE
 PUBMED 8798544
 REFERENCE 2 (bases 1 to 2088)
 AUTHORS Briskin,M.J. and Shyjan,A.M.
 TITLE Direct Submission
 JOURNAL Submitted (24-MAY-1996) LeukoSite Inc., 215 First Street,
 Cambridge, MA 02142, USA

FEATURES

source
 1..2088
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 36..1772
 /codon_start=1
 /product="hyaluronan synthase"
 /protein_id="AAC50706.1"
 /db_xref="GI:1556465"
 /translation="MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAGV
 PLASDRYGLLAFGLYGLFSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTIS
 AYQDPRYLRLQCLASARALYPRARLVVYDGNRAEDLYMDFREVFADDEPATY
 VWDGNVHPWEPAAGAGAGAVEAEDEPGLRVEALVRTRCVCAVQWGGKREV
 MYTFKALGSDVDYVQVCDSDTRDPMALVELVLDPRVAGVGGVDRILNPJDSW
 VSFLLSLRYVAFNVACQSFHCVSCISGSLGLYRNLIQQELKAWYNQFLGTHC
 TFGDRLHNLNMGVATKTSRSRCVSETPFLRLWSLQTRNSKYSFRELWYNAL
 WWHRHAWMTYEAVVSGLFPFPFVALRLFLYAGRPWALLWVLLCVQVVALKAFAA
 WLRCRLRLVLSLYAPLYMCLPAKFLALVTMNSQGWGTSGRKLAANYVPLPLAL
 WALLLLGLVRSVAHEARADWSGSPRAAEAYHLAAGAGYVGVYVWVAMLTLYVWVGVRRL

ORIGIN		CRRRTGGYRVQV"	
Query Match		86.1%; Score 933; DB 9; Length 2088;	
Best Local Similarity		90.6%; Pred. No. 3.6e-114;	
Matches 1048; Conservative		0; Mismatches 35; Indels 74; Gaps 2;	
Qy	1	ATGAGACAGCAGGAGCGCCCAAGCCCACTCTGACGCGCCGCGCTGCTCGGGCTGGCC	60
Db	36	ATGAGACAGCAGGAGCGCCCAAGCCCACTCTGACGCGCCGCGCTGCTCGGGCTGGCC	95
Qy	61	CGGAGGGTGTGACCATCGCCTTGGCCCTGCTCATCTGCGGCTCATGACCTGGCCCTTAC	120
Db	96	CGGAGGGTGTGACCATCGCCTTGGCCCTGCTCATCTGCGGCTCATGACCTGGCCCTTAC	155
Qy	121	GCCGCGGGGTGCGCTGCGCTGCGATGCTGACGCGCTCTGCGCTTGGGCTTACGCG	180
Db	156	GCCGCGGGGTGCGCTGCGCTGCGATGCTGACGCGCTCTGCGCTTGGGCTTACGCG	215
Qy	181	GCCTTCTTTTACGCGCAGCTGGTGGGCGAGAGCTCTTTCGGGTACCTTGGAGACCGCGCG	240
Db	216	GCCTTCTTTTACGCGCAGCTGGTGGGCGAGAGCTCTTTCGGGTACCTTGGAGACCGCGCG	275
Qy	241	GTGCGCGCGCGCGCGCGCGCGCGCTGGATGCGAGCCACCGCGCGCTGTGGGCTGACC	300
Db	276	GTGCGCGCGCGCGCGCGCGCGCGCTGGATGCGAGCCACCGCGCGCTGTGGGCTGACC	335
Qy	301	ATCTCGCGCTTACGAGGAGACCCGCGCTACCTGCGCAGTGCCTTGGGCTTGGGCTGACC	360
Db	336	ATCTCGCGCTTACGAGGAGACCCGCGGTACTGGGCGAGTGCCTTGGGCTTGGGCTGACC	395
Qy	361	CTGCTGTACCCGCGCGCGCGCTGCGGCTCTCATGTGTGTGATGCAACCGCGCGGAG	420
Db	396	CTGCTGTACCCGCGCGCGCGCTGCGGCTCTCATGTGTGTGATGCAACCGCGCGGAG	455
Qy	421	GACCTCTACATGTGTGACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCCACCTAC	480
Db	456	GACCTCTACATGTGTGACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCCACCTAC	515
Qy	481	GTGTGGGAGCGGCAACTACCAACGAGCCCTGGGAAACCCCGCGCGCGCGGTGGGCGCC	540
ORIGIN		Query Match 86.1%; Score 933; DB 6; Length 2116;	
Best Local Similarity		90.6%; Pred. No. 3.6e-114;	
Matches 1048; Conservative		0; Mismatches 35; Indels 74; Gaps 2;	
Qy	1	ATGAGACAGCAGGAGCGCCCAAGCCCACTCTGCGAGCCGCGCTGCTCGGCTGGCC	60
Db	36	ATGAGACAGCAGGAGCGCGCCCAAGCCCACTCTGCGAGCCGCGCTGCTCGGCTGGCC	95
Qy	61	CGGAGGGTGTGACCATCGCCTTGGCCCTGCTCATCTGCGGCTCATGACCTGGGCTTAC	120
Db	96	CGGAGGGTGTGACCATCGCCTTGGCCCTGCTCATCTGCGGCTCATGACCTGGGCTTAC	155
Qy	121	GCCGCGGGGTGCGCTGCGCTTCCGATCGCTACGCGCTCTGCGCTTTCGGGCTTACGCG	180
Db	156	GCCGCGGGGTGCGCTGCGCTTCCGATCGCTACGCGCTCTGCGCTTTCGGGCTTACGCG	215
Qy	181	GCCTTCTTTTACGCGCAGCTGGTGGGCGAGAGCTCTTTCGGGTACCTTGGAGACCGCGCG	240
Db	216	GCCTTCTTTTACGCGCAGCTGGTGGGCGAGAGCTCTTTCGGGTACCTTGGAGACCGCGCG	275
Qy	241	GTGCGCGCGCGCGCGCGCGCGCTGGATGCGAGCCACCGCGCGCTGTGGGCTGACC	300
Db	276	GTGCGCGCGCGCGCGCGCGCGCTGGATGCGAGCCACCGCGCGCTGTGGGCTGACC	335
Qy	301	ATCTCGCGCTTACGAGGAGACCCGCGCTACCTGCGCAGTGCCTTGGGCTTGGGCTGACC	360
Db	336	ATCTCGCGCTTACGAGGAGACCCGCGGTACTGGGCGAGTGCCTTGGGCTTGGGCTGACC	395
Qy	361	CTGCTGTACCCGCGCGCGCTGCGGCTCTCATGTGTGTGATGCAACCGCGCGGAG	420
Db	396	CTGCTGTACCCGCGCGCGCTGCGGCTCTCATGTGTGTGATGCAACCGCGCGGAG	455
Qy	421	GACCTCTACATGTGTGACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCCACCTAC	480
Db	456	GACCTCTACATGTGTGACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCCACCTAC	515
Qy	481	GTGTGGGAGCGGCAACTACCAACGAGCCCTGGGAAACCCCGCGCGCGGTGGGCGCC	540
ORIGIN		Query Match 86.1%; Score 933; DB 6; Length 2116;	
Best Local Similarity		90.6%; Pred. No. 3.6e-114;	
Matches 1048; Conservative		0; Mismatches 35; Indels 74; Gaps 2;	
Qy	1	ATGAGACAGCAGGAGCGCCCAAGCCCACTCTGCGAGCCGCGCTGCTCGGCTGGCC	60
Db	36	ATGAGACAGCAGGAGCGCGCCCAAGCCCACTCTGCGAGCCGCGCTGCTCGGCTGGCC	95
Qy	61	CGGAGGGTGTGACCATCGCCTTGGCCCTGCTCATCTGCGGCTCATGACCTGGGCTTAC	120
Db	96	CGGAGGGTGTGACCATCGCCTTGGCCCTGCTCATCTGCGGCTCATGACCTGGGCTTAC	155
Qy	121	GCCGCGGGGTGCGCTGCGCTTCCGATCGCTACGCGCTCTGCGCTTTCGGGCTTACGCG	180
Db	156	GCCGCGGGGTGCGCTGCGCTTCCGATCGCTACGCGCTCTGCGCTTTCGGGCTTACGCG	215
Qy	181	GCCTTCTTTTACGCGCAGCTGGTGGGCGAGAGCTCTTTCGGGTACCTTGGAGACCGCGCG	240
Db	216	GCCTTCTTTTACGCGCAGCTGGTGGGCGAGAGCTCTTTCGGGTACCTTGGAGACCGCGCG	275
Qy	241	GTGCGCGCGCGCGCGCGCGCGCTGGATGCGAGCCACCGCGCGCTGTGGGCTGACC	300
Db	276	GTGCGCGCGCGCGCGCGCGCGCTGGATGCGAGCCACCGCGCGCTGTGGGCTGACC	335
Qy	301	ATCTCGCGCTTACGAGGAGACCCGCGCTACCTGCGCAGTGCCTTGGGCTTGGGCTGACC	360
Db	336	ATCTCGCGCTTACGAGGAGACCCGCGGTACTGGGCGAGTGCCTTGGGCTTGGGCTGACC	395
Qy	361	CTGCTGTACCCGCGCGCGCTGCGGCTCTCATGTGTGTGATGCAACCGCGCGGAG	420
Db	396	CTGCTGTACCCGCGCGCGCTGCGGCTCTCATGTGTGTGATGCAACCGCGCGGAG	455
Qy	421	GACCTCTACATGTGTGACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCCACCTAC	480
Db	456	GACCTCTACATGTGTGACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCCACCTAC	515
Qy	481	GTGTGGGAGCGGCAACTACCAACGAGCCCTGGGAAACCCCGCGCGCGGTGGGCGCC	540
Db	516	GTGTGGGAGCGCAACTACCAACGAGCCCTGGGAAACCCCGCGCGCGGTGGGCGCC	575
Qy	541	GGAGCCTATCGGAGGTGGAGGCGGAGGATCTTGGGCGGCTTGGCAGTGGAGGCGCTGGTG	600
Db	576	GGAGCCTATCGGAGGTGGAGGCGGAGGATCTTGGGCGGCTTGGCAGTGGAGGCGCTGGTG	635
Qy	601	AGGACTCGCAGTGCCTGCTGGCGCAGCGCTTGGGCGGCAAGCGCGAGTCACTGTAC	660
Db	636	AGGACTCGCAGTGCCTGCTGGCGCAGCGCTTGGGCGGCAAGCGCGAGTCACTGTAC	695
Qy	661	ACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTCAGAGTCTTGTGACTCGGACACA	720
Db	696	ACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTCAGAGTCTTGTGACTCGGACACA	755
Qy	721	AGGTTGAGCCCATGCGCAGCTCGTGGGCTGCTGAGGCTTGGAGAGGACCCCGCGGTA	780
Db	756	AGGTTGAGCCCATGCGCAGCTCGTGGGCTGCTGAGGCTTGGAGAGGACCCCGCGGTA	815
Qy	781	GGGCGCTGTGGTGGGAGCTGCGGATCTTAACCTCTGGAATCTTGGGCTGAGTCTTCTTA	840
Db	816	GGGCGCTGTGGTGGGAGCTGCGGATCTTAACCTCTGGAATCTTGGGCTGAGTCTTCTTA	875
Qy	841	AGCAGCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGACTACTTCCAC	900
Db	876	AGCAGCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGACTACTTCCAC	935
Qy	901	TGTGTATCTGTCATCAGGGTCTCTAG-----	928
Db	936	TGTGTATCTGTCATCAGGGTCTCTAGGGCTCTATAGGCTATATAGGAATAACCTTTCGAGCAGTTT	995
Qy	929	-----AATCTGCGCCAGGCGCCAGGAGCAGCGAGTATGCC	965

ALLLGLVRSVAHEARDWSGPRAAAEVHLAAGAGAYGVYVWVAMLTLYWVGVRRLC
RRRTGYRVQV"

ORIGIN

Query Match 85.0%; Score 920.8; DB 9; Length 2119;
Best Local Similarity 90.3%; Pred. No. 1.5e-112;
Matches 1037; Conservative 0; Mismatches 37; Indels 74; Gaps 2;

Qy 10 CAGAGCGGCCAAGCCACTCTTCGACGCCCGCGCTGCTCGGCTGCGCGGAGGGTG 69
Db 48 CAGAGCGGCCAAGCCACTCTTCGACGCCCGCGCTGCTCGGCTGCGCGGAGGGTG 107
Qy 70 CTGACCATCGCTTTCGCGCTCTCATCTCGGCGCTCATGACCTGCGCTGCGCGG 129
Db 108 CTGACCATCGCTTTCGCGCTCTCATCTCGGCGCTCATGACCTGCGCTGCGCGG 167
Qy 130 GTGCGCTGCGCTTCGATCGCTACGCGCTCTGCGCTTCGCGCTCTACGCGGCTTCCTT 189
Db 168 GTGCGCTGCGCTTCGATCGCTACGCGCTCTGCGCTTCGCGCTTCGCGGCTTCCTT 227
Qy 190 TCAGGCACTGCTGCGCAGAGCTCTTCGCTACCTGGAGCACCGCGGCTGCGCGG 249
Db 228 TCAGGCACTGCTGCGCAGAGCTCTTCGCTACCTGGAGCACCGCGGCTGCGCGG 287
Qy 250 GCGCGCGCGCGCTGCGATGCGACACCGCGCGCAGTGTGCGCTGACCATCTCGGC 309
Db 288 GCGCGCGCGCGCTGCGATGCGACACCGCGCGCAGTGTGCGCTGACCATCTCGGC 347
Qy 310 TACGAGGAGACCCCGCTACTCTGCGCAGTGTGCGCTGCGCGCGCGCTCTGTAC 369
Db 348 TACGAGGAGACCCCGCTACTCTGCGCAGTGTGCGCTGCGCGCGCGCTCTGTAC 407
Qy 370 CCGCGCGCGCGCTGCGGCTCTCATGCTGTGATGCGCAACCGCGCGGAGCTCTAC 429
Db 408 CCGCGCGCGCGCTGCGGCTCTCATGCTGTGATGCGCAACCGCGCGGAGCTCTAC 467
Qy 430 ATGCTCGACATGTTCCGCGAGTCTTCGCTCAGCAGGACCCCGCAGCTAGCTGTGGAC 489
Db 468 ATGCTCGACATGTTCCGCGAGTCTTCGCTCAGCAGGACCCCGCAGCTAGCTGTGGAC 527
Qy 490 GGCAACTACCAACGACCTTGGAAACCCCGCGCGCGCGCGCTGCGCGCGGAGCTAT 549
Db 528 GGCAACTACCAACGACCTTGGAAACCCCGCGCGCGCGCGCTGCGCGCGGAGCTAT 587
Qy 550 CCGAGGTGGAGCGGAGGATCTTGGGCGCTGCGAGTGGAGCGCTGTTGAGACTCGC 609
Db 588 CCGAGGTGGAGCGGAGGATCTTGGGCGCTGCGAGTGGAGCGCTGTTGAGACTCGT 647
Qy 610 AGTGGCTGTGCTGGCGCAGCGCTGGGCGGCAAGCGCGAGTCTGTTACACAGCTTC 669
Db 648 AGTGGCTGTGCTGGCGCAGCGCTGGGCGGCAAGCGCGAGTCTGTTACACAGCTTC 707
Qy 670 AAGCGCTCGGAGATTCGGTGAATACGTGAGGTCTGTGATCGGACACAAAGTTGGAC 729
Db 708 AAGCGCTCGGAGATTCGGTGAATACGTGAGGTCTGTGATCGGACACAAAGTTGGAC 767
Qy 730 CCGATGCACTGCTGGAGCTGTCGGGTACTGACGAGGACCCCGGCTAGGGCTGTT 789
Db 768 CCGATGCACTGCTGGAGCTGTCGGGTACTGACGAGGACCCCGGCTAGGGCTGTT 827
Qy 790 GGTGGGACGTGCGGATCTTAACCTCTGAGCTTCGCGGTACGTTCTTAAGCAGCTG 849
Db 828 GGTGGGACGTGCGGATCTTAACCTCTGAGCTTCGCGGTACGTTCTTAAGCAGCTG 887
Qy 850 CGATCTCGGCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCACTGTGTATCC 909
Db 888 CGATCTCGGCTTCAATGTGGAGCGGCTTGTGAGAGCTACTTCCACTGTGTATCC 947
Qy 910 TGCATCAGCGGCTCTCTAG----- 928
Db 948 TGCATCAGCGGCTCTCTAGGAGCTATAGGAATAACCTTTGACGAGTCTTTTGGGCC 1007
Qy 929 -----AATCCTGCCAGGCCCGGAGCAGCGGATGATGCCCTCATTCCT 974

Db 1008 TGGTACAACCAAGAGTTCTCTGGGTACCCACTGTACTTTTGGGATGACCGGCACCTCAC 1067
Qy 975 CGCCCC-----CGTGAGGTACACCTCCAGGTCCGGTCTGTACTCA 1015
Db 1068 AACCGCATGCTCAGCATGGGTATGCTACCAAGTACACCTCCAGGTCCGGTCTGTACTCA 1127
Qy 1016 GAGAGCGCTTCTCTCTGCGGTGCTGAGCAGCAGACAGCTGCTCAAGTCTGTAC 1075
Db 1128 GAGAGCGCTTCTCTCTGCGGTGCTGAGCAGCAGACAGCTGCTCAAGTCTGTAC 1187
Qy 1076 TTCCGTCA 1083
Db 1188 TTCCGTCA 1195

RESULT 4

HUMAS LOCUS HUMAS 2108 bp mRNA linear PRI 06-FEB-1999
DEFINITION Homo sapiens mRNA for hyaluronan synthase, complete cds.

ACCESSION D84424

VERSION D84424.1 GI:1401033

KEYWORDS hHAS; hyaluronan synthase.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2108)

AUTHORS Itano, N. and Kimata, K.

TITLE Molecular cloning of human hyaluronan synthase

JOURNAL Biochem. Biophys. Res. Commun. 222 (3), 816-820 (1996)

MEDLINE 96244584

PUBMED 8651928

REFERENCE 2 (bases 1 to 2108)

AUTHORS Itano, N.

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 2108)

AUTHORS Itano, N.

TITLE Direct Submission

Submitted (16-APR-1996) Naoki Itano, Aichi Medical University,
Institute for Molecular Science of Medicine, Nagakute, Aichi-gun,
Aichi 480-11, Japan (Tel:052-264-4811(ex.2087), Fax:0561-63-3532)

FEATURES

source

1..2108 /organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="brain"

/clone_lib="igt11"

/dev_stage="fetal"

1..2108 /genes="hHAS"

149..1780 /genes="hHAS"

/product="hyaluronan synthase"

/protein_id="BA012351.1"

/db_xref="GI:1401034"

/translation="MTWYAAGVPLASDRYGLLAGFLYSAHLVAQSLFAYLEHR

RVAAGALPDATARSVALTISAYQEDPAYLRQCLASARALLYPRARVRLMVVDGN

RAEDLYVMDFREVFADEDPATYVNDGNTHQWPEAAGAVGAGYREVEADPGRLA

VEALVTRCVCVAORWGKREVMYTAFAKLGDSVDYVOVCDSDRLDPMALLVRLV

LDEDPRVAGGVDRILNPLDSWVSFLSLRYMVAFNVERACQSPHYCVSCI SGPLGL

YRNLLQQLFEAWINOKFLGTHCTFGDDRLHRLMSMGYATKTSRSRCYSETPSPF

LRWLSQOTRWSKYFRENLYNALMWRHHAWMTYEAVSGLFFPFAATVLRLLFYAGR

PMALLVLLCQGVALLAKAFAWLRGLRMLLSLYAPLYMCGLLPAKFLALVTNQ

SGMGTSGRKLAAVYVPLLPALWALLLGLVRSVAHEARADWSGPRAAEYHLAA

GAGAYGVYVWVAMLTLYWVGVRRLCRRRTGGRVQV"

ORIGIN

Query Match 85.0%; Score 920.6; DB 9; Length 2108;
Best Local Similarity 90.2%; Pred. No. 1.5e-112;
Matches 1038; Conservative 0; Mismatches 39; Indels 74; Gaps 2;

Qy 7 CAGCAGGACGGCCCAAGCCCACTCTCTGAGCCCGCGCTGCTCGGCTTGGCCCGGAGG 66
Db CAGCAGGACGGCCCAAGCCCACTCTCTGAGCCCGCGCTGCTCGGCTTGGCCCGGAGG 109
Qy 67 GTGCTGACATCGCTTCCCTTCCCTGCTCATCTGAGCCCTCATGACCTGGCCCTACCGCC 126
Db GTGCTGACATCGCTTCCCTTCCCTGCTCATCTGAGCCCTCATGACCTGGCCCTACCGCC 169
Qy 127 GGGGTGCGCTGGCTTCCGCTACGCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 186
Db GGGGTGCGCTGGCTTCCGCTACGCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 229
Qy 187 CTTTCAGCGCACCTGGTGGCGAGAGCTTCTGCGTACCTTGAGACACCGCGGGTGGCG 246
Db CTTTCAGCGCACCTGGTGGCGAGAGCTTCTGCGTACCTTGAGACACCGCGGGTGGCG 289
Qy 247 GCGGCGCGCGGGGCGCTGGATGACGACCGCGCGCGAGTGGCGCTGACCATCTCC 306
Db GCGGCGCGCGGGGCGCTGGATGACGACCGCGCGCGAGTGGCGCTGACCATCTCC 349
Qy 307 GCCTACCAAGGAGACCGCGTACCTGCGCAGTCCCTGGGCTCCGCGCGCGCTGCTG 366
Db GCCTACCAAGGAGACCGCGTACCTGCGCAGTCCCTGGGCTCCGCGCGCGCTGCTG 409
Qy 367 TACCGCGCGCGCGCTTCCCTCATGTGGTGGTGGCAACCGCGCGCGAGACCTC 426
Db TACCGCGCGCGCGCTTCCCTCATGTGGTGGTGGCAACCGCGCGCGAGACCTC 469
Qy 427 TACATGTGCGATGTTCCGCGAGTCTTCCGCTGAGAGGACCCCGCGCGAGTGG 486
Db TACATGTGCGATGTTCCGCGAGTCTTCCGCTGAGAGGACCCCGCGCGAGTGG 529
Qy 487 GACCGCAACTACCAACGAGCCCTGGGAAACCGCGCGCGCGGGTGGCGCGCGAGCC 546
Db GACCGCAACTACCAACGAGCCCTGGGAAACCGCGCGCGCGGGTGGCGCGCGAGCC 589
Qy 547 TATCGGAGGTGGAGCGCGAGGATCTTGGGGGCTGGCAGTGGAGGCGCTGGTAGGACT 606
Db TATCGGAGGTGGAGCGCGAGGATCTTGGGGGCTGGCAGTGGAGGCGCTGGTAGGACT 649
Qy 607 GCGAGGTGCGTGGTGGCGAGCGCTGGGGCGCGCAAGCGCGAGTCTATGACAGCC 666
Db GCGAGGTGCGTGGTGGCGAGCGCTGGGGCGCGCAAGCGCGAGTCTATGACAGCC 709
Qy 667 TTCAAGCGCTCGAGATTCGCTGAGTACCTGAGGCTCTGCTCGGACACAGGTTG 726
Db TTCAAGCGCTCGAGATTCGCTGAGTACCTGAGGCTCTGCTCGGACACAGGTTG 769
Qy 727 GACCCCATGGCACTGCTGGGCTGCTGAGGAGGACCCCGCGGTAGGGCT 786
Db GACCCCATGGCACTGCTGGGCTGCTGAGGAGGACCCCGCGGTAGGGCT 829
Qy 787 GTTGTGGGACGTGGGATCTTAAACCTCTGAGCTCTGGGTAGGTTCTTAAGCAGC 846
Db GTTGTGGGATGTGGGATCTTAAACCTCTGAGCTCTGGGTAGGTTCTTAAGCAGC 889
Qy 847 CTGCGATCTGGGTAGGCTTCAATGTGGAGCGGGTGTGAGGCTACTTCCACTGTGTA 906
Db CTGCGATCTGGGTAGGCTTCAATGTGGAGCGGGTGTGAGGCTACTTCCACTGTGTA 949
Qy 907 TCCTGCATCAGCGGTCTCTTAG----- 928
Db TCCTGCATCAGCGGTCTCTTAG----- 1009
Qy 929 -----AATCTGCCAGGCCCCGAGGAGCAGCGATGATGCCCTCAT 971
Db GCCTGGTACAACAGAGTCTCTGGGTACCACCTGTACTTTTGGGATGACCGGACCTC 1069
Qy 972 CCGCGCC-----CGTGCAGGTACACCTCGAGTCCGCTGCTAC 1012
Db ACCAACCGCATGCTCAGCATGGGTATATGCTACCAAGTACACCTCAGGTCCTGCTGCTAC 1129

Qy 1013 TCAGAGACGCGCTCTGCTCTCTGCGGTGGCTGAGCCAGCAGACACGCTGCTCAAGTCG 1072
Db TCAGAGACGCGCTCTGCTCTCTGCGGTGGCTGAGCCAGCAGACACGCTGCTCAAGTCG 1189
Qy 1073 TACTTCCGTGA 1083
Db TACTTCCGTGA 1200
RESULT 5
AR137038
LOCUS AR137038 2117 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6162908.
ACCESSION AR137038
VERSION AR137038.1 GI:14478288
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2117)
AUTHORS Itano,N. and Kimata,K.
TITLE Polypeptide of human-origin hyaluronate synthetase and DNA encoding the same
JOURNAL Patent: US 6162908-A 1 19-DEC-2000;
FEATURES Location/Qualifiers
source 1..2117
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 85.0%; Score 920.6; DB 6; Length 2117;
Best Local Similarity 90.2%; Pred. No. 1.5e-112;
Matches 1038; Conservative 0; Mismatches 39; Indels 74; Gaps 2;
Qy 7 CAGCAGGACGGCCCAAGCCCACTCTCTGAGCCCGCGCTGCTCGGCTTGGCCCGGAGG 66
Db 50 CAGCAGGACGGCCCAAGCCCACTCTCTGAGCCCGCGCTGCTCGGCTTGGCCCGGAGG 109
Qy 67 GTGCTGACATCGCTTCCCTTCCCTGCTCATCTGAGCCCTCATGACCTGGCCCTACCGCCGC 126
Db 110 GTGCTGACATCGCTTCCCTTCCCTGCTCATCTGAGCCCTCATGACCTGGCCCTACCGCCGC 169
Qy 127 GGGGTGCGCTGGCTTCCGCTACGCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 186
Db 170 GGGGTGCGCTGGCTTCCGCTACGCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCC 229
Qy 187 CTTTCAGCGCACCTGGTGGCGAGAGCTTCTGCGTACCTTGAGACACCGCGGGTGGCG 246
Db 230 CTTTCAGCGCACCTGGTGGCGAGAGCTTCTGCGTACCTTGAGACACCGCGGGTGGCG 289
Qy 247 GCGGCGCGCGGGGCGCTGGATGACGACCGCGCGCGAGTGGCGCTGACCATCTCC 306
Db 290 GCGGCGCGCGGGGCGCTGGATGACGACCGCGCGCGAGTGGCGCTGACCATCTCC 349
Qy 307 GCCTACCAAGGAGACCCCGCTACCTGCGCAGTCCCTGGGCTCCGCGCGCGCTGCTG 366
Db 350 GCCTACCAAGGAGACCCCGCTACCTGCGCAGTCCCTGGGCTCCGCGCGCGCTGCTG 409
Qy 367 TACCGCGCGCGCGCTTCCCTCATGTGGTGGTGGCAACCGCGCGCGAGACCTC 426
Db 410 TACCGCGCGCGCGCTTCCCTCATGTGGTGGTGGCAACCGCGCGCGAGACCTC 469
Qy 427 TACATGTGCGATGTTCCGCGAGTCTTCCGCTGAGAGGACCCCGCGCGAGTGG 486
Db 470 TACATGTGCGATGTTCCGCGAGTCTTCCGCTGAGAGGACCCCGCGCGAGTGG 529
Qy 487 GACCGCAACTACCAACGAGCCCTGGGAAACCGCGCGCGCGGGTGGCGCGCGAGCC 546
Db 530 GACCGCAACTACCAACGAGCCCTGGGAAACCGCGCGCGCGGGTGGCGCGCGAGCC 589
Qy 547 TATCGGAGGTGGAGCGCGAGGATCTTGGGGGCTGGCAGTGGAGGCGCTGGTAGGACT 606
Db 590 TATCGGAGGTGGAGCGCGAGGATCTTGGGGGCTGGCAGTGGAGGCGCTGGTAGGACT 649

Qy	607	CGCAGGTCGGTGTCTGGCGCACGCCTGGGGCGCAAGCGCGAGGTCTATGTACACAGCC	666
Db	650	CGCAGGTCGGTGTCTGGCGCACGCCTGGGGCGCAAGCGCGAGGTCTATGTACACAGCC	709
Qy	667	TTCGAAGCGGTCGAGATTGGTGGAATACTGTGAGGTCTGTGACTGGAACAAGTTG	726
Db	710	TTCGAAGCGGTCGAGATTGGTGGAATACTGTGAGGTCTGTGACTGGAACAAGTTG	769
Qy	727	GACCCCATGCACCTGCTGGAGCTCGTGGGGTA CTGGACGAGGACCCCCGGGTAGGGGCT	786
Db	770	GACCCCATGGACACTGCTGGAGCTCGTGGGGTA CTGGACGAGGACCCCCGGGTAGGGGCT	829
Qy	787	GTTGGTGGGACAGTGGCGGATCCCTTAAACCCTCTGGACTCCTGGGTGAGCTTCTTAAGCAGC	846
Db	830	GTTGGTGGGATGTGGGATCCCTTAAACCCTCTGGACTCCTGGGTGAGCTTCTTAAGCAGC	889
Qy	847	CTGCGAATACTGGGTAGCTTCAATGTGGAGGGGCTGTGACAGACTACTTCCACTGTGTA	906
Db	890	CTGCGAATACTGGGTAGCTTCAATGTGGAGGGGCTGTGACAGACTACTTCCACTGTGTA	949
Qy	907	TCCTGCATCAGCGTCTCTAG-----	928
Db	950	TCCTGCATCAGCGTCTCTAGGCTATATAGGAATAACCTCTTGGACAGTTCTTCTTGAG	1009
Qy	929	-----AATCTGCCACGAGATTCCTGGGTACCCACTGTACTTTTGGGATGACCGCACCTC	971
Db	1010	GCCTGGTACAACAGAGATTCCTGGGTACCCACTGTACTTTTGGGATGACCGCACCTC	1069
Qy	972	CCTCGCCCC-----CGTGCAGGTACACCTCCAGGTCCCCTGCTGCTAC	1012
Db	1070	ACCAACCGCATGCTCAGCATGGGTATGTCTACCAAGTACACCTCCAGGTCCCCTGCTGCTAC	1129
Qy	1013	TCAGAGAGCGCCTCGTCTTCTCGGTGGTGTAGCCAGCAGACACGCTGGTCCAAGTCG	1072
Db	1130	TCAGAGAGCGCCTCGTCTTCTCGGTGGTGTAGCCAGCAGACACGCTGGTCCAAGTCG	1189
Qy	1073	TACTTCCGTGA 1083	
Db	1190	TACTTCCGTGA 1200	
RESULT 6			
AY463695		2109 bp mRNA linear PRI 07-DEC-2003	
LOCUS	Papio anubis hyaluronan synthase (HAS1) mRNA, complete cds.		
DEFINITION	Papio anubis hyaluronan synthase (HAS1) mRNA, complete cds.		
ACCESSION	AY463695		
VERSION	AY463695.1 GI:38607341		
KEYWORDS			
SOURCE	Papio anubis (olive baboon)		
ORGANISM	Papio anubis		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;		
TITLE	Cercopithecinae; Papio.		
JOURNAL	1 (bases 1 to 2109)		
REFERENCE	Martinez-Duncker, I., Oriol, R. and Mollicone, R.		
AUTHORS	Evolution of the hyaluronan, nodulation c, chitin and cellulose		
TITLE	synthases: a superfamily of cell-wall associated carbohydrate		
JOURNAL	polymerizing enzymes		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 2109)		
TITLE	Martinez-Duncker, I., Oriol, R. and Mollicone, R.		
JOURNAL	Direct Submission		
FEATURES	Submitted (12-NOV-2003) Unite 504 Glycobiology Et Signalisation		
source	Cellulaire, Inserm, 16 Avenue Paul Vaillant Couturier, Villejuif		
gene	94807, France		
	Location/Qualifiers		
	1..2109		
	/organism="Papio anubis"		
	/mol_type="mRNA"		
	/db_xref="taxon:9555"		
	1..2109		
	/gene="HAS1"		

QY 706 TGTGACTCGGACACAGCTTGACCCCATGGACCTGCTGGAGCTCGTGGGGTACTGGAC 765
Db 756 TGTGACTCGGACACAGCTTGACCCCATGGACCTGCTGGAGCTCGTGGGGTACTGGAC 815
QY 766 GAGACACCCCGGTAGGGGCTGTGGTGGGACGTCGGATCCTTAACCCCTCGACTCC 825
Db 816 GAGACACCCCGGTAGGGGCTGTGGTGGGACGTCGGATCCTTAACCCCTCGACTCC 875
QY 826 TGGGTGAGCTTCTTAAGCAGCTCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 885
Db 876 TGGGTGAGCTTCTTAAGCAGCTCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 935
QY 886 CAGAGCTACTTCCACTGTGATCTCTGATCAGCGTCTCTCTAG 928
Db 936 CAGAGCTACTTCCACTGTGATCTCTGATCAGCGTCTCTCTAG 978

RESULT 7

AC018755/c
LOCUS AC018755 229155 bp DNA linear PRI 26-JUL-2000
DEFINITION Homo sapiens chromosome 19, BAC BC330783 (CIT-HSPC_470E3), complete
sequence.

AC018755

AC018755.3 GI:9454515

VERSION HTG.

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 229155) Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 2 (bases 1 to 229155) Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 3 (bases 1 to 229155) Lamerdin, J.E., McCready, P.M., Skowronski, E., Viswanathan, V.,

REFERENCE 4 (bases 1 to 229155) Burkhardt-Schultz, K., Gordon, L., Dias, J., Ramirez, M., Stilwag, S.,

REFERENCE 5 (bases 1 to 229155) Phan, H., Velasco, N., Do, L., Regala, W., Terry, A., Brower, A.,

REFERENCE 6 (bases 1 to 229155) Barnes, J., Danganan, L., Erler, A., Christensen, M., Georgescu, A.,

REFERENCE 7 (bases 1 to 229155) Avila, J., Liu, S., Andrise, T., Frankheim, M., Attix, C.,

REFERENCE 8 (bases 1 to 229155) Amico-Keller, G., Coefield, J., Duarte, S., Lucas, S., Bruce, R.,

REFERENCE 9 (bases 1 to 229155) Thomas, P., Quan, G., Kronmiller, B., Arellano, A., Sanders, C., Ow, D.,

REFERENCE 10 (bases 1 to 229155) Nolan, M., Trong, S., Kobayashi, A., Olsen, A.S. and Carrano, A.V.

REFERENCE 11 (bases 1 to 229155) Sequence analysis of a 5-Mb region in 19q13.4

REFERENCE 12 (bases 1 to 229155) Unpublished

REFERENCE 13 (bases 1 to 229155) DOE Joint Genome Institute.

REFERENCE 14 (bases 1 to 229155) Direct Submission

REFERENCE 15 (bases 1 to 229155) Submitted (19-DEC-1999) Production Sequencing Facility, DOE Joint

REFERENCE 16 (bases 1 to 229155) Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 17 (bases 1 to 229155) Lamerdin, J.E.

REFERENCE 18 (bases 1 to 229155) Direct Submission

REFERENCE 19 (bases 1 to 229155) Submitted (26-JUL-2000) DOE Joint Genome Institute, Lawrence

REFERENCE 20 (bases 1 to 229155) Livermore National Laboratory, Livermore, CA 94550

REFERENCE 21 (bases 1 to 229155) On Jul 26, 2000 this sequence version replaced gi:7458780.

REFERENCE 22 (bases 1 to 229155) Map and sequence oriented from centromere to q-telomere. BC330783

REFERENCE 23 (bases 1 to 229155) (CIT-HSPC_470E3) is overlapped on the left by BC849408

REFERENCE 24 (bases 1 to 229155) (CITB-EL_3073N11, AC020914) from bases 1 to 125 of this accession,

REFERENCE 25 (bases 1 to 229155) and overlaps cosmid R28782 (LNLN-R_248F10, AC005946) on the right

REFERENCE 26 (bases 1 to 229155) from bases 217,905 to 229,155 of this accession. Additional

REFERENCE 27 (bases 1 to 229155) chromosome 19 map and sequence information may be obtained at:

REFERENCE 28 (bases 1 to 229155) http://www.bio.lnl.gov/bbrp/genome/genome.html.

FEATURES

source

1. .229155

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="19"

/map="19q13.4"

/clone="CIT-470E3 (LNLN clone name BC330783)"

/clone_lib="CalTech human BAC library C"

201. .262

/rpt_family="L2"

<299_16870

/gene="ZNF175"

/note="zinc finger protein 175; synonym: OTK18"

join(<299. .550,8540. .8666,8936. .9031,13776. .16870)

CDS

/gene="ZNF175"
/product="ZNF175 mRNA"
/join(479. .550,8540. .8666,8936. .9031,13776. .15616)
/gene="ZNF175"
/codon_start=1
/product="zinc finger protein 175"
/protein_id="AAF87843.1"
/db_xref="GI:9454517"
/translation="MPADVNLSSQKQPVLPGLGQDSCBASVFEDVTVDPSREEQWL
DPAQCLYRDVMLELYSHLFAVGYHINPEVI FRLMEKEPRVEAEVSHQRCQREF
GLEIPQKEISKASFOQDMVGEPTRDGSCWCSILELRDLADRTKKDEQNOIQPMHSA
FFNKKTLDNNSCYKDPGKMRIRPHLASSQKQPKCCLFTESLKLNLVNGQNESN
DTEQLDDVWGSGQLFSSSDACSKNTHGETFCGKQCRVKQGLSKLQHQIHQK
KPDGSCGGSFTQKSHLFAQQRHSHVGNLHCGKCGKAFQMLKSLYLDHTDIP
CICKEGKVFIORSELLTHQTRKKPKYKCHDCGKAFQMLSLFHQTHSREKLYE
CSECGKGFSONSLI IHQKIHGTGEROVACSECGKAFQKSLSLHQRHSGOKSVCI
ECGQAFIQKALHIVHQSHTGEKPYQCHNCKSFISQSLDIHRIHTGERPYECSDC
GKTPTQKSHLNHQRHHTGERHHVCSGCKAFNOKSILSHQRIHTGERPKYKCGKAF
AFTSKSQFQKHQRIHTGERPYVTECGKAFNRSNFKHQIHTHTRPFPVYKCGKAF
VQKSELTHQTHMGKPYECLDCGKSFSSKPKQLKQVHQRHHTGERPYVCSGCKAFN
RSNFKHQIHTHTRDKSYKCSYKVGFTKQ"
1097. 1198
/rpt_family="MER99"
1559. .1829
/rpt_family="AluJb"
complement(2130. .2355)
/rpt_family="L1ME"
complement(2359. .2627)
/rpt_family="AluY"
complement(2816. .3183)
/rpt_family="HAL1"
complement(3290. .3598)
/rpt_family="AluSg"
complement(3611. .3760)
/rpt_family="MIR"
complement(3766. .4067)
/rpt_family="HAL1"
4098. .4215
/rpt_family="AluJb"
complement(4339. .4406)
/rpt_family="HAL1"
complement(4518. .4724)
/rpt_family="HAL1"
complement(4803. .5110)
/rpt_family="AluY"
5266. .5561
/rpt_family="AluSg"
5635. .5932
/rpt_family="AluJo"
5984. .6120
/rpt_family="AluJb"
6121. .6423
/rpt_family="AluSx"
6424. .6579
/rpt_family="AluJb"
6647. .6913
/rpt_family="L1ME"
6936. .7295
/rpt_family="MER33"
7700. .7785
/rpt_family="L2"
complement(9216. .9434)
/rpt_family="L1ME"
complement(9435. .9745)
/rpt_family="AluY"
complement(9746. .9815)
/rpt_family="L1ME"
complement(10009. .10303)
/rpt_family="AluSp"
11783. .12637
/rpt_family="L1ME2"
12180. .12330
/gene="ZNF175"
STS

/standard_name="SHGC-32377"
13436. .15619
/rpt_family="MIR"
16068. .16112
/rpt_family="(TA)n"
complement(16113. .16329)
/rpt_family="MER46A"
16330. .16390
/rpt_family="(TA)n"
16403. .16439
/rpt_family="(TATG)n"
16793. .17013
/rpt_family="MERS8A"
17024. .17052
/rpt_family="(CA)n"
17311. .17545
/rpt_family="MIR"
complement(18683. .18896)
/rpt_family="MIR"
complement(18915. .19183)
/rpt_family="AluY"
19336. .19599
/rpt_family="MIR"
19811. .19928
/rpt_family="GA-rich"
22181. .23081
/rpt_family="L2"
complement(23692. .23901)
/rpt_family="L2"
complement(23902. .24201)
/rpt_family="AluSg1"
complement(24202. .24651)
/rpt_family="L2"
26136. .26434
/rpt_family="LTR12"
26435. .26525
/rpt_family="LTR30"
26526. .26643
/rpt_family="LTR12"
complement(26686. .26884)
/rpt_family="AluSg/x"
27065. .27240
/rpt_family="MIR"
complement(27238. .27554)
/rpt_family="L2"
28406. .28433
/rpt_family="GC-rich"
28641. .28678
/rpt_family="GC-rich"
complement(28997. .29259)
/rpt_family="MSR1"
complement(30219. .30489)
/rpt_family="L2"
30753. .30882
/rpt_family="FLAM_C"
31138. .31416
/rpt_family="AluY"
complement(31547. .31716)
/rpt_family="MER3"
32041. .32112
/rpt_family="L1MEC"
32117. .32516
/rpt_family="MLT2E"
35056. .35118
/rpt_family="MIR"
complement(35578. .35636)
/rpt_family="MERSA"
complement(35690. .35820)
/rpt_family="MIR"
35872. .35902
/rpt_family="(TG)n"
complement(35906. .36209)
/rpt_family="AluJo"

Query Match	64.4%	Score 697.2	DB 9	Length 229155
Best Local Similarity	99.6%	Pred. No. 1.2e-83	Indels 0	Gaps 0
Matches 699	Conservative 0	Mismatches 3		
Qy 7	CAGCAGGACGCGCCCAAGCCCACTCTCTGACAGCCCGCGCTGCTCCGGCTGCGCCGAGG 66			
Db 147050	CAGCAGGACGCGCCCAAGCCCACTCTCTGACAGCCCGCGCTGCTCCGGCTGCGCCGAGG 146991			
Qy 67	GTGCTGACCATGCGCTTTCGCTGCTCATCTGCGCTCTCATAGACTGGGCTTACGCGCC 126			
Db 146990	GTGCTGACCATGCGCTTTCGCTGCTCATCTGCGCTCTCATAGACTGGGCTTACGCGCC 146931			
Qy 127	GSGGTGCGCTGCGCTCCGATCGCTAGCGGCTCTGCGCTTCCGCTTACGGGGCCTTC 186			
Db 146930	GSGGTGCGCTGCGCTCCGATCGCTAGCGGCTCTGCGCTTCCGCTTACGGGGCCTTC 146871			
Qy 187	CTTTACGCGCACCTGTGTGGCGCAGAGCTCTTTCGCTTACCTGGAGCACCGCGCGCTGCG 246			
Db 146870	CTTTACGCGCACCTGTGTGGCGCAGAGCTCTTTCGCTTACCTGGAGCACCGCGCGCTGCG 146811			
Qy 247	GCGCGCGCGCGCGCGCTTGTGATGACCCACCGCGCGCAGTGTGGCTGACCATCTCC 306			
Db 146810	GCGCGCGCGCGCGCGCTTGTGATGACCCACCGCGCGCAGTGTGGCTGACCATCTCC 146751			
Qy 307	GCCTTACGAGGAGGACCCCGGCTACCTGCGCGAGTGTGCGCTCCGCGCGCTGCTG 366			
Db 146750	GCCTTACGAGGAGGACCCCGGCTACCTGCGCGAGTGTGCGCTCCGCGCGCTGCTG 146691			
Qy 367	TACCGCGCGCGCGCTGCGCTCTCATGTGTGATGGCAACCGCGCGCGAGGACTC 426			
Db 146690	TACCGCGCGCGCGCTGCGCTCTCATGTGTGATGGCAACCGCGCGCGAGGACTC 146631			
Qy 427	TACATGTCGACATGTTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCGCTGCTG 486			
Db 146630	TACATGTCGACATGTTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCGCTGCTG 146571			
Qy 487	GAGCGCAACTACCACGAGCCCTGGGACCCCGCGCGCGCGCGCGCGCGCGAGCC 546			
Db 146570	GAGCGCAACTACCACGAGCCCTGGGACCCCGCGCGCGCGCGCGCGCGAGCC 146511			
Qy 547	TATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGGCGCTGGTGGAGCT 606			
Db 146510	TATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGGCGCTGGTGGAGCT 146451			
Qy 607	CGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCTATGACAGCC 666			
Db 146450	CGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCTATGACAGCC 146391			
Qy 667	TTCAAGCGCTCGGAGATTTCGGTGGACTACGTGCGGTCTGT 708			
Db 146390	TTCAAGCGCTCGGAGATTTCGGTGGACTACGTGCGGTGAGT 146349			
RESULT 8				
AB097568		2095 bp	mrna	linear
LOCUS	AB097568			ROD 28-APR-2004
DEFINITION	Rattus norvegicus HAS1 mRNA for hyaluronan synthase 1, complete cds.			
ACCESSION	AB097568			
VERSION	AB097568.1	GI:26453348		
KEYWORDS				
SOURCE	Rattus norvegicus (Norway rat)			
ORGANISM	Rattus norvegicus			
REFERENCE	1			
AUTHORS	Icano,N., Sawai,T., Atsumi,F., Miyaishi,O., Taniguchi,S., Kannagi,R., Hamaguchi,M. and Kinata,K.			
TITLE	Selective expression and functional characteristics of three Mammalian hyaluronan synthases in oncogenic malignant transformation			

J. Biol. Chem. 279 (18), 18679-18687 (2004)
14724275
REFERENCE
2 (bases 1 to 2095)
Itano,N.
Direct Submission
Submitted (05-DEC-2002) Naoki Itano, Aichi Medical University, Aichi,
Institute for Molecular Science of Medicine, Nagakute, Aichi, Aichi,
480-1195, Japan (E-mail: itano@amugw.aichi-med-u.ac.jp,
Tel:81-52-264-4811 (ex.2095), Fax:81-561-63-3532)
Location/Qualifiers
1..2095
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
1..2095
/gene="HAS1"
43..1794
/gene="HAS1"
/codon_start=1
/product="hyaluronan synthase 1"
/protein_id="BAC43730.1"
/db_xref="GI:26453349"
/translation="NRQMPKPESEARCCSGIARRVLITITFALLILGLMTWAYAGVP
LASDYGGLALGLYGLAFSLVAQSLFAYLEHRRVTVAAARFAPKAGFADRAATARSVA
LTISAYQDPTLYRQCLTSARALLYPRLTFLVMDVGNRAEDFLVMDVDFREVFAD
PATVYMDGNTHQWPEPAEAGAVEGAYREVEAEDPRLAVALVTRVCVCAQWVG
GKREVMYAFKALGSDVVOVCDSDRLDPALDELVRVLDDEPRVAGVGDVRLN
PLDSWVFLSLRYVWA FNVERACOSYFHCVSCI SGPLGLYRNLLQOFLNAYNOKF
LGTHTCTFDDHRLNRLMNGYATKYTSRSCYSETSPSFLNLSQQTRWKSYPREW
LYNALWHRHAAWMTYEAVVSGLFFPFAATVRLFYAGRPWALLWLLVCQGVALAK
AFAAWLRGCLRLLSLVPLMCGLLPAPKFLALVTWNGWGTSGRKLAAANYVP
LPLMALLLGLLIRSVQAEVRADWSGFSRAAEAYHLAAGASAYVAYVWIMLTIVYV
GVRLCRRRSGYRVQV"

ORIGIN
Query Match 62.7%; Score 679.4; DB 10; Length 2095;
Best Local Similarity 84.6%; Pred. No. 1.1e-80;
Matches 793; Conservative 0; Mismatches 126; Indels 18; Gaps 2;
QY 10 CAGGACGGCCCAAGCCACTCTCCAGCCGCGCTCTCGGCTCGCCGGAGGTG 69
DB 49 CAGGACATCCCAAGCCCTCAGAGCAGCAGCTTCTCGCTCGGCTCGCCAGCGGTG 108
QY 70 CTGACCATCGCTTCGCCCTGCTCATCTCGGCTCATGACCTGGGCTTACGCCCGCGG 129
DB 109 CTCACGATCACCTTCGCCCTGCTCATCTCGGCTCATGACCTGGGCTTACGCCAGGA 168
QY 130 GTGCCGCTGGCTTCGATCGCTACGGCTCTCTGGCTTCTCGGCTTACGGGCTTCTCTT 189
DB 169 GTACCTCTGGCTTCTGATCCCTATGGCTCTCTGGCTTCTGGCTTCTATGGGCGTTCCTC 228
QY 190 TCAGGACACCTGGTGGCGCAGACCTCTTCGGTACCTGGAGCAGCGCGGCTGGCGG 249
DB 229 AGTGGCACCTAGTGGCAGAGCCTCTTCGGCTTACCTGGAGCAGCGAGGCTGACCGGTG 288
QY 250 GCGGCGCG-----GCGGCGCTGGATGACGACCGCGCGGCTGGCGG 294
DB 289 GCTGGCGCGCGCTTTTTCGAGAGGACCCCTGGATGGGCACTGGCGGAGCGTGGCA 348
QY 295 CTGACCATCTCGCTTACGAGAGGACCCCGCTTACCTGGCGCAGTGGCTGGCGTCCGCC 354
DB 349 CTCACCATCTCAGCCTACGAGGAGACCCACTTACCTGGCGCAGTGGCTTACCTCCGCG 408
QY 355 CGCGCCCTCTGATACCGCGCGCTGGCTGGCTTCTCATGGTGGTGGTGGTGGTGGTGGTGG 414
DB 409 CGCGCTTCTGCTGATACCGCGCAGCGGCTGGCGCTGCTTATGGTGGTGGTGGTGGTGG 468
QY 415 GCCGAGGACCTCTACATGCTGACATGTTCCGCGAGGCTTCTCGCTGACGAGGACCCCGCC 474
DB 469 GCGGAGGATCTGATGCTGGACATGTTCCGAGAGTCTTCGAGATGAGGACCTCGCC 528
QY 475 ACGTACGTGTGGGACGGCACTTACCAACGAGCCCTGGGAACCGCC---GGCGCGCGCGCG 531
|||||

Db 529 ACTTATGTGTGGATGGCAACTACCATCAGCCTTGGGAGCCAGCTGAGCGCGCGGTGCT 588
QY 532 GTGGCGCGCGGAGCCTATCGGAGGTGAGGCGGAGGATCCTGGCGGCTGGCAGTGGAG 591
Db 599 GTGGGTGAAGTGCCTACCGCGAGGTGAGGCTGAGGACCTTGGCGGCTGGCGGTAGAG 648
QY 592 GCGCTGTGTGAGGACTCGCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 651
Db 649 GCGCTGTGTGAGGACCCCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 708
QY 652 GTCTATGTACACAGCCTTCAAGCGCTCGGAGGATTCGCTGCTGCTGCTGCTGCTGCTGCTG 711
Db 709 GTCTATGTACACAGCCTTCAAGCGCTCGGAGGATTCGCTGCTGCTGCTGCTGCTGCTGCTG 768
QY 712 TCGGACACAAAGGTGGACCCCATGCTGCTGAGCTCGTGGCGGTACTGACGAGGAGAC 771
Db 769 TCAGACACAAAGGTGAGACCCCATGCTGCTGAGCTTGTGCGAGTGTGCTGATGAGAC 828
QY 772 CCGCGGTGAGGGCTGTTGGTGGGACCTGCGGATCTTAAACCTCTTGAATCTCTGGGTG 831
Db 829 CCGCGGTGAGGTGCTGTTGGAGGAGATGTGAGAACTCTTAAACCTCTTGGACTCTTGGGTG 888
QY 832 AGCTTCTTAAAGCAGCTCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 891
Db 889 AGCTTCTTAAAGCAGCTCGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948
QY 892 TACTTCCACTGTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 928
Db 949 TACTTCCACTGTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 985

RESULT 9
E13681
LOCUS
E13681
DEFINITION
DNA encoding novel mouse hyaluronate synthetase.
ACCESSION
E13681.1
VERSION
GI:3252450
KEYWORDS
JP 1997224674-A/1.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 2102)
AUTHORS
Itano,N. and Kimata,H.
TITLE
POLYPEPTIDE OF NEW HYALURONIC ACID-SYNTHETIC ENZYME AND DNA CODING
THE SAME
JOURNAL
Patent: JP 1997224674-A 1 02-SEP-1997;
SCIENCE & TECH AGENCY
OS Mus musculus (mouse)
PN JP 1997224674-A/1
PD 02-SEP-1997
PF 26-FEB-1996 JP 1996038336
PI ITANO NAKKI, KIMATA HIROHARU
PC C12N15/09,C07H21/04,C07K14/47,C12N9/00//A61K48/00,C12N1/21, PC
(C12N9/00,
PC C12R1:91),(C12N9/00,C12R1:19),(C12N1/21,C12R1:19); CC
strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key Location/Qualifiers
FT source 1..2102
FT /organism="Mus musculus"
FT /cell_line="FM3A"
FT 5'UTR 1..48
FT CDS 49..1800
FT /product="Novel mouse hyaluronate synthetase"
FT 3'UTR 1801..2102
FT Location/Qualifiers
1..2102
/organism="Mus musculus"
/mol_type="genomic DNA"


```
ORIGIN
/db_xref="taxon:10090"

Query Match 62.6%; Score 677.8; DB 6; Length 2102;
Best Local Similarity 84.5%; Pred. No. 1.8e-80;
Matches 792; Conservative 0; Mismatches 127; Indels 18; Gaps 2;

Qy 10 CAGAGCGCGCCCAAGCCACTCCTCGACGCGCGCGCTGCTCGGCGCTGCGCGGAGGGTG 69
Db 55 CAGGACATGCCAAAGCCCTCAGAGCAGCGGGTGTCTCTCGGCGCTGCGCGGAGGCA 114
Qy 70 CTGACCATCGCTTCGCGCTGCTCATCTCGGCGCTCATGACTCGGCGCTCATGCGCGCGG 129
Db 115 CTCAGATCATCTTGGCGCTGCTCATCTCGGCGCTCATGACTCGGCGCTCATGCGCGG 174
Qy 130 GTGCGCTGCGCTCGATCGCTACGCGCTCTCGGCGCTCTGCGCGCTTACGCGGCGCTTCTT 189
Db 175 GTTCTCTGCGCTTCAGATCGCTATGGACTCTGCGCGCTTGGCGCTTATGGGCGATTCTC 234
Qy 190 TCAGCGCACCTGTGCGCGCAGAGCCTCTTGGCGCTTACGCGCTTACGCGCGGCGGCG 249
Db 235 AGCGCACACCTAGTGGCGCAGAGCCTCTTGGCGCTTACGCGCTTACGCGCGGCGGCG 294
Qy 250 GCGGCGCG-----GGGCGCGCTGGATGCGCAGCCAGCGCGCGAGTGTGGCG 294
Db 295 GCTGCGGCGCGCTCTTGGGCGAGGGCGCCCTGGATGCGCGCTTACGCGCGGCGTGGCA 354
Qy 295 CTGACCATCTCGCGCTTACCGCGGAGGACCCCGCGTACCTGCGCGCAGTGCTGCGCGTCCGCC 354
Db 355 CTCACCATCTCAGCGCTACCAAGAGGATCCCGCTTACCTGCGCGCAGTGCTTACCTCGCG 414
Qy 355 CGGCGCTTGTATACCGCGCGCGCGCTGCGCGCTCTCATGTGTGTGGATGGCAACCGC 414
Db 415 CGGCGCTTGTATACCGCGCAGCGAGTTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 474
Qy 415 GCGGAGGACCTCTACATGTTGCGATGTTGCGAGGTTCTTGGTGTGAGGAGCCCGCGC 474
Db 475 GCTGAGGATCTGTATATGTTGGAATGTTTGGAGAGTCTTGGCGGATGAGGAGCCCGCGC 534
Qy 475 ACGTATGTGGGAGCGGCAACTACACCGCGCTTGGGAGCCCGCGCGGCGGCGGCGGCGG 531
Db 535 ACTTATGTGGGATGGCAACTACCATCAGCGCTGGGAAACAGCGAGGCTTACGCGGCGCT 594
Qy 532 GTGGGCGCGGAGCCTATCGGAGGTGAGCGGAGGATCTTGGCGCGCTGCGCGCTGCGAGTGGAG 591
Db 595 GTCGCTGAAGGTGCTTACCGGAGGTGAGCGGAGGACCCCGCGCGGCTTGGCGGCTGAG 654
Qy 592 GCGCTGTGAGGACTCGCAGGTGCGTGTGCGCGCAGCGCTGCGGCGCGCAAGCGCGAG 651
Db 655 GCGCTGTGAGAACACGCGAGGTGCGTGTGCGTGTGCGTGTGCGTGTGCGGCGCGCAAGCTGAG 714
Qy 652 GTCATGTACACAGCGCTTCAAGCGCTCGGAGATTCGGTGTGACTAGTGTGAGTCTGTGAC 711
Db 715 GTCATGTACACAGCTTTCAAGGCACTGGGCGACTCCGTTGGACTACGTGCGAGTCTGTGAC 774
Qy 712 TCGGACACAGGTTGGAACCCCATGGCACTGTGAGCTGTGCGGCTGTGAGCAGGAG 771
Db 775 TCAGACACAGACTAGAACCCCATGGCACTGTGAGCTTGTGGAGCTTGTGGAGTGTGGATGAAGAC 834
Qy 772 CCGCGGTGAGGGCTGTTGGTGGGAGCGTGGCGGATCCCTTAACCTCTGGAATCTCGGGTC 831
Db 835 CCGCGGTGAGGGCTGTTGGAGGGGATGTGAGGATCTCTTAACCTCTGGAATCTCGGGTC 894
Qy 832 AGCTTCTAAGCAGCTGCGATATCGGATAGCCTTCAATGTGGAGCGCGGCTTGTGAGAGC 891
Db 895 AGCTTCTGAGCAGCTTTCGATATCGGATAGCCTTCAATGTGGAAACGAGCTTGTGAGAGC 954
Qy 892 TACTTCCACTGTGTATCTTGCATCAGCGGCTCTCTAG 928
Db 955 TACTTCCACTGTGTATCTTGCATCAGTGGTCTCTCTGG 991

FEATURES
source
1..2102
/organism="Mus sp."
/mol_type="genomic DNA"
/db_xref="taxon:10095"

ORIGIN

Query Match 62.6%; Score 677.8; DB 6; Length 2102;
Best Local Similarity 84.5%; Pred. No. 1.8e-80;
Matches 792; Conservative 0; Mismatches 127; Indels 18; Gaps 2;

Qy 10 CAGAGCGCGCCCAAGCCACTCCTCGACGCGCGCGCTGCTCGGCGCTGCGCGGAGGGTG 69
Db 55 CAGGACATGCCAAAGCCCTCAGAGCAGCGGGTGTCTCTCGGCGCTGCGCGGAGGCA 114
Qy 70 CTGACCATCGCTTCGCGCTGCTCATCTCGGCGCTCATGACTCGGCGCTCATGCGCGG 129
Db 115 CTCAGATCATCTTGGCGCTGCTCATCTCGGCGCTCATGACTCGGCGCTCATGCGCGG 174
Qy 130 GTGCGCTGCGCTCGATCGCTACGCGCTCTCGGCGCTCTGCGCGCTTACGCGGCGCTTCTT 189
Db 175 GTTCTCTGCGCTTCAGATCGCTATGGACTCTGCGCGCTTGGCGCTTATGGGCGATTCTC 234
Qy 190 TCAGCGCACCTGTGCGCGCAGAGCCTCTTGGCGCTTACGCGCTTACGCGCGGCGGCG 249
Db 235 AGCGCACACCTAGTGGCGCAGAGCCTCTTGGCGCTTACGCGCTTACGCGCGGCGGCG 294
Qy 250 GCGGCGCG-----GGGCGCGCTGGATGCGCAGCCAGCGCGCGAGTGTGGCG 294
Db 295 GCTGCGGCGCGCTCTTGGGCGAGGGCGCCCTGGATGCGCGCTTACGCGCGGCGTGGCA 354
Qy 295 CTGACCATCTCGCGCTTACCGCGGAGGACCCCGCGTACCTGCGCGCAGTGCTGCGCGTCCGCC 354
Db 355 CTCACCATCTCAGCGCTACCAAGAGGATCCCGCTTACCTGCGCGCAGTGCTTACCTCGCG 414
Qy 355 CGGCGCTTGTATACCGCGCGCGCGCTGCGCGCTCTCATGTGTGTGGATGGCAACCGC 414
Db 415 CGGCGCTTGTATACCGCGCAGCGAGTTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 474
Qy 415 GCGGAGGACCTCTACATGTTGCGATGTTGCGAGGTTCTTGGTGTGAGGAGCCCGCGC 474
Db 475 GCTGAGGATCTGTATATGTTGGAATGTTTGGAGAGTCTTGGCGGATGAGGAGCCCGCGC 534
Qy 475 ACGTATGTGGGAGCGGCAACTACACCGCGCTTGGGAGCCCGCGCGGCGGCGGCGGCGG 531
Db 535 ACTTATGTGGGATGGCAACTACCATCAGCGCTGGGAAACAGCGAGGCTTACGCGGCGCT 594
Qy 532 GTGGGCGCGGAGCCTATCGGAGGTGAGCGGAGGATCTTGGCGCGCTGCGCGCTGCGAGTGGAG 591
Db 595 GTCGCTGAAGGTGCTTACCGGAGGTGAGCGGAGGACCCCGCGCGGCTTGGCGGCTGAG 654
Qy 592 GCGCTGTGAGGACTCGCAGGTGCGTGTGCGCGCAGCGCTGCGGCGCGCAAGCGCGAG 651
Db 655 GCGCTGTGAGAACACGCGAGGTGCGTGTGCGTGTGCGTGTGCGTGTGCGGCGCGCAAGCTGAG 714
Qy 652 GTCATGTACACAGCGCTTCAAGCGCTCGGAGATTCGGTGTGACTAGTGTGAGTCTGTGAC 711
Db 715 GTCATGTACACAGCTTTCAAGGCACTGGGCGACTCCGTTGGACTACGTGCGAGTCTGTGAC 774
Qy 712 TCGGACACAGGTTGGAACCCCATGGCACTGTGAGCTGTGCGGCTGTGAGCAGGAG 771
Db 775 TCAGACACAGACTAGAACCCCATGGCACTGTGAGCTTGTGGAGCTTGTGGAGTGTGGATGAAGAC 834
Qy 772 CCGCGGTGAGGGCTGTTGGTGGGAGCGTGGCGGATCCCTTAACCTCTGGAATCTCGGGTC 831
Db 835 CCGCGGTGAGGGCTGTTGGAGGGGATGTGAGGATCTCTTAACCTCTGGAATCTCGGGTC 894
Qy 832 AGCTTCTAAGCAGCTGCGATATCGGATAGCCTTCAATGTGGAGCGCGGCTTGTGAGAGC 891
Db 895 AGCTTCTGAGCAGCTTTCGATATCGGATAGCCTTCAATGTGGAAACGAGCTTGTGAGAGC 954
Qy 892 TACTTCCACTGTGTATCTTGCATCAGCGGCTCTCTAG 928
Db 955 TACTTCCACTGTGTATCTTGCATCAGTGGTCTCTCTGG 991

RESULT 10
```


Matches 792; Conservative 0; Mismatches 127; Indels 18; Gaps 2;

QY 10 CAGGACGCGCCCAAGCCACTCTCTGAGCCCGCGCTGCTCGGCGCTGCGCGGAGGCTG 69
Db |||||
55 CAGGACATGCCAAAGCCCTCAGAGCAGCGCTGCTGCTCTGCGCTGCGCAGCGCAGCA 114
QY 70 CTGACATCGCTTGGCCCTGCTCATCTCTGGCGCTCATGACCTGGCGCTCATCGCGCGCGG 129
Db |||||
115 CTCAGATCATCTTGGCCCTGCTCATCTCTGGCGCTCATGACCTGGCGCTCATCGCGCAGGC 174
QY 130 GTGCGCGCTGCGCTCGATCGCTACCGCTCTCTGGCGCTCTGGCGCTCTATGGCGCATCTCTC 189
Db |||||
175 GTCTCTCTGGCTTCAGATCGCTATGACCTCTCTGGCGCTCTGGCGCTCTATGGCGCATCTCTC 234
QY 190 TCAGCGCACCTGCTGGCGCAGAGCTCTCTCGGCTACCTGGAGCACCGCGCGCTGGCGCG 249
Db |||||
235 AGCGCACACTAGTGGCACAGAGCTCTCTCGCTTACCTGGAGCACCGAGGCTGGCAGCG 294
QY 250 GCGCGCGC-----GGGCGCGCTGATGACAGCCACCGCGCGCTGATGGCG 294
Db |||||
295 GCTGCGCGCGCTCTCTTGGCGAAGGCGCGCTCTGATGCGGCGCTGACGACGCTGGCA 354
QY 295 CTGACATCTCGCTTACAGAGACCGCGCTGCTGCGCGCTGCTGCGCTGCTGCGCGC 354
Db |||||
355 CTCACATCTCAGCTTACCAAGAGATCCCGCTTACCTGCGCGCTGCTGACCTCCGCG 414
QY 355 CGCGCGCTGCTGACCGCGCGCGCTGCGCTCTCATGCTGCTGATGCGCAACCGC 414
Db |||||
415 CGCGCTGCTGCTGACCGCACAGAGTTACGCTGCTCATGCTGCTGAGCGCAACCGC 474
QY 415 GCGGAGGACCTTACATGCTGCATGTTTCCGCGAGTCTTCCGCTGACGAGGACCCCGC 474
Db |||||
475 GCTGAGGATCTTACATGCTGACATGTTCCGAGAAGTCTTCCGCGATGAGGACCCCGC 534
QY 475 AGCTAGCTGCGAGCGGCACTACACAGCCCTGCGAACCCTGCGCG-----CGGCGG 531
Db |||||
535 ACTTATGTTGGGATGGCACTTACCATCAGCCCTGGGAACCGAGCGGCTACGCGCGCT 594
QY 532 GTGGCGCGCGGCTTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGCGAGTGGAG 591
Db |||||
595 GTGCGTGAAGTGTCTTACCGGAGGTGGAGCGCGAGGACCCCGCGGCTTGGCGTGGAG 654
QY 592 GCGCTGTTGAGGACTCGCAGGTGCGTGTGCGGCGCAGCGCTGGGCGCGCAGCGCAG 651
Db |||||
655 GCGCTGTTGAGAACCGCAGGTGCGTGTGCGTGGCTCAGCGTGGCGCGCAACGCTGAG 714
QY 652 GTCATGTACAGCCCTTCAAGCGCTCGAGATTCGCTGGACTACGTGAGGTCTGTGAC 711
Db |||||
715 GTCATGTACAGCTTCAAGGCACTGGGCGGACTCGTGGGACTACGTGAGGTCTGTGAC 774
QY 712 TCGGACACAGGTGGACCCCATGCACTGCTGGAGCTGCTGGGCTACTGGACGAGGAC 771
Db |||||
775 TCAGACACAAGACTAGACCCCATGAGCACTGCTGGAGCTGTGCGAGTGTGATGAAGAC 834
QY 772 CCGCGGTAGGGGCTGTGCTGGGAGCGTGGGATCTTAAACCTCTTGAACCTCTGGGTC 831
Db |||||
835 CCGCGGTAGGGGCTGTGAGGAGGATGTGAGGATCTTAAACCTCTTGAACCTCTGGGTC 894
QY 832 AGCTTCTTAAGCAGCTGCGATCTAGGCTGAGCTTCAATGTGGAGCGGCTGTGAGAGC 891
Db |||||
895 AGCTTCTTGAAGCAGCTTCTGATCTAGGCTGAGCTTCAATGTGGACGAGCTGTGAGAGC 954
QY 892 TACTTCCATGTATCTCTGATCAGCGCTCTCTAG 928
Db |||||
955 TACTTCCATGTGTGCTGCTGATCAGTGTGCTCTCTAG 991

RESULT 14

AC130783/c
LOCUS AC130783 193986 bp DNA linear HTG 19-NOV-2002
DEFINITION Pan troglodytes clone CH251-426A12, WORKING DRAFT SEQUENCE, 8
ordered pieces.
ACCESSION AC130783

VERSION
KEYWORDS
SOURCE
ORGANISM

AC130783.2 GI:25100968
HTG; HTGS PHASE2; HTGS DRAFT.
Pan troglodytes (chimpanzee)
Pan troglodytes

REFERENCE
AUTHORS

1 (bases 1 to 193986)
Akter N., Antonellis A., Ayele K., Beckstrom-Sternberg S.M.,
Benjamin B., Blakesley R.W., Bouffard G.G., Brinkley C., Brooks S.,
Cariaga K., Coleman B., Engle J., Granite S., Guan X., Gupta J.,
Hachichi P., Han J., Hansen N., Ho S.-L., Idol J.R., Karlins E.,
Laric P., Lee-Lin S.-Q., Legaapi R., Maduro Q.L., McDowell J.,
Margulies E.H., Masello C., Maskeri B., Prasad A.,
Paguigan C., Pearson R., Portnoy M.E., Prasad A.,
Reddix-Dugue N., Schandler K., Schueler M.G., Sison C.,
Stratipod S., Thomas J.W., Thomas P.J., Touchman J.W., Vogt J.L.,
Weberby K.D., Wiggins L., Young A. and Green E.D.
NISC Comparative Sequencing Initiative

TITLE
JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 193986)
Green E.D.
Direct Submission

TITLE
JOURNAL

Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
Grovmont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 193986)
Green E.D.
Direct Submission

REFERENCE
AUTHORS

Submitted (19-NOV-2002) NIH Intramural Sequencing Center, 8717
Grovmont Circle, Gaithersburg, MD 20877, USA
On Nov 19, 2002 this sequence version replaced gi:22218453.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoehngri.nih.gov
----- Project Information
Center project name: dxd
Center clone name: 426A12

COMMENT

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 191764 bases at least Q40
Consensus quality: 192510 bases at least Q30
Consensus quality: 192998 bases at least Q20
Insert size: 184000; agarose-fp
Insert size: 193286; sum-of-contigs
Quality coverage: 11.49x in Q20 bases; agarose-fp
Quality coverage: 10.94x in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 11512: contig of 11512 bp in length

```
* 11513 11612: gap of unknown length
* 59488: contig of 47876 bp in length
* 59489 103266: contig of 43678 bp in length
* 103267 127885: contig of 24519 bp in length
* 127886 129885: contig of unknown length
* 129886 129892: contig of 1707 bp in length
* 129893 129792: gap of unknown length
* 129793 131077: contig of 1285 bp in length
* 131078 174663: contig of 43486 bp in length
* 174664 174763: gap of unknown length
* 174764 193986: contig of 19223 bp in length.
FEATURES
    source
        1..193986
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            /clone="CH251-426A12"
            /clone_lib="CH251"
        1..11512
            /note="assembly_fragment"
            clone_end:8p6
            vector_side:left
        1..76
            /note="clone overlaps with GenBank Accession Number
            AC132069 clone CH251-132A6 (center project name dxc)"
            11613..59488
                /note="assembly_fragment"
            59489..103266
                /note="assembly_fragment"
            103267..127885
                /note="assembly_fragment"
            127886..129885
                /note="assembly_fragment"
            129886..131077
                /note="assembly_fragment"
            131078..174663
                /note="assembly_fragment"
            174664..193986
                /note="assembly_fragment"
                clone_end:T7
                vector_side:right

ORIGIN
Query Match 61.8%; Score 669; DB 2; Length 193986;
Best Local Similarity 97.2%; Pred. No. 6.6e-80;
Matches 697; Conservative 0; Mismatches 5; Indels 15; Gaps 1;

Qy 7 CAGCAGGACGCGCCAGCCACCTCTCTGACGCCCGCGCTGCTCGCGCTGCGCGGAGG 132769
Db 132828 CAGCAGGACGCGCCAGCCACCTCTCTGACGCCCGCGCTGCTCGCGCTGCGCGGAGG 132769

Qy 67 GTGCTGACCATCGCTTCCGCTGCTATCTGCGCTCATGAGCTGCGCTACGCGCC 126
Db 132768 GTGCTGACCATCGCTTCCGCTGCTATCTGCGCTCATGAGCTGCGCTACGCGCC 132709

Qy 127 GGGGTGCGCTGCGCTCGATCGCTACGCGCTCTGCGCTTCCGCTTACGCGCGCTTC 186
Db 132708 GGGGTGCGCTGCGCTCGATCGCTACGCGCTCTGCGCTTCCGCTTACGCGCGCTTC 132649

Qy 187 CTTTCAGCGCACCTGCTGCGCAGAGCCTCTTTCGCTACCTGAGACACCGCGCGGTGCG 246
Db 132648 CTTTCAGCGCACCTGCTGCGCAGAGCCTCTTTCGCTACCTGAGACACCGCGCGGTGCG 132589

Qy 247 -----GCGCGCGCGCGCGCGCGCGCTGATGACGACCGCGCGCAGTGTG 291
Db 132588 GCGCGCGCGCGCGCGCGCGCGCGCGCGCTGATGACGACCGCGCGCAGTGTG 132529

Qy 292 GCGCTGACCATCTCCGCCCTACAGAGAGACCCCGGTACCTGCGCCAGTGTGCTGCGCTCC 351
Db 132528 GCGCTGACCATCTCCGCCCTACAGAGAGACCCCGGTACCTGCGCCAGTGTGCTGCGCTCC 132469

Qy 352 GCCCGCGCTGCTGTACCCGCGCGCGCGCTGCGCTCTCATGCTGCTGATGGCAAC 411
Db 132468 GCCCGCGCTGCTGTGTACCCGCGCGCGCGCTGCGCTCTCATGCTGCTGATGGCAAC 132409

Qy 412 CGCGCCGAGGACCTCTACATGCTCGACATGTTCCGCGAGGCTCTTCGCTGACGAGGACCCC 471
Db 132408 CGCGCCGAGGACCTCTACATGCTCGACATGTTCCGCGAGGCTCTTCGCTGACGAGGACCCC 132349

Qy 472 GCCACGTAGCTGTGGGACGGCAACTACACAGCCCTGGGAAACCCGCGCGCGCGCGCGCG 531
Db 132348 GCCACGTAGCTGTGGGACGGCAACTACACAGCCCTGGGAAACCCGCGCGCGCGCGCGCG 132289

Qy 532 GTGGCGCGCGGAGCCTATCGGAGGTGAGCGGAGGATCTTGGCGCGCTGCGCAGTGTGAG 591
Db 132288 GTGGCGCGCGGAGCCTATCGGAGGTGAGCGGAGGATCTTGGCGCGCTGCGCAGTGTGAG 132229

Qy 592 GCGCTGCTGAGGACTCGCAGGTGCTGTGCTGCGCGAGCGCTGCGGCGCGCAAGCGCGAG 651
Db 132228 GCGCTGCTGAGGACTCGCAGGTGCTGTGCTGCGCGAGCGCTGCGGCGCGCAAGCGCGAG 132169

Qy 652 GTCATGTACACAGCCTTCAAGCGCGCTCGGAGATTGCGTGACACTACGTGCGAGTCTGT 708
Db 132168 GTCATGTACACAGCCTTCAAGCGCGCTCGGAGATTGCGTGACACTACGTGCGAGTCTGT 132112

RESULT 15
AC137058/c
LOCUS
DEFINITION Papio anubis clone RP41-126M5, WORKING DRAFT SEQUENCE, 10 ordered
pieces.
AC137058
VERSION AC137058.3 GI:28460766
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Papio anubis (olive baboon)
ORGANISM Papio anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopitheciinae; Papio.
REFERENCE 1 (bases 1 to 185623)
AUTHORS Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.I., McDowell,J.,
Margulies,E.H., Masiello,C., Maskeri,B., Prasad,A.,
Paguirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
REFERENCE 2 (bases 1 to 185623)
AUTHORS Green,E.D.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
REFERENCE 3 (bases 1 to 185623)
AUTHORS Green,E.D.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2003) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
On Feb 21, 2003 this sequence version replaced gi:27476124.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@hgri.nih.gov
----- Project Information
Center project name: dtj
Center clone name: 126M05

The sequence data in this record represents an 'enhanced'
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 01:47:21 ; Search time 703.833 Seconds
(without alignments)
9108.809 Million cell updates/sec

Title: US-10-672-399-5
Perfect score: 1083
Sequence: 1 atgagacagcagcgcc.....tccaagtcgtactccgtga 1083

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002as:*
7: Geneseq2002bs:*
8: Geneseq2003as:*
9: Geneseq2003bs:*
10: Geneseq2003cs:*
11: Geneseq2003ds:*
12: Geneseq2004as:*
13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	933	85.1	1737	10 ADL13691	Adl13691 Osteoarth
2	933	85.1	2088	10 ADL13690	Adl13690 Osteoarth
3	933	85.1	2116	2 AAT99541	Aat99541 Human hya
4	931.4	85.0	1737	10 ADL13692	Adl13692 Osteoarth
5	931.4	85.0	2087	10 ADL13694	Adl13694 Osteoarth
6	911	84.1	2117	2 AAT96713	Aat96713 Human hya
7	697.2	64.4	23122	10 ADL13693	Adl13693 Osteoarth
8	677.8	62.6	1752	8 ABZ76734	Abz76734 Mouse hya
9	677.8	62.6	1752	10 AAD59442	Aad59442 Mouse hya
10	677.8	62.6	2102	2 AAT91855	Aat91855 Mouse hya
11	677.8	62.6	2102	3 AAZ88199	Aaz88199 Mouse hya
12	677.8	62.6	2102	3 AAZ88199	Aaz88199 Mouse hya
13	677.8	62.6	2102	3 AAA39987	Aaa39987 Murine HA
14	422	39.0	662	6 ABQ27384	Abq27384 Oligonuc
15	422	39.0	662	6 ABQ27385	Abq27385 Oligonuc
16	393.8	36.4	662	6 ABQ27382	Abq27382 Oligonuc
17	393.8	36.4	662	6 ABQ27383	Abq27383 Oligonuc
18	333.2	31.3	490	9 ACH39998	Ach39998 Human toe
19	290.8	26.9	1783	10 ADC49212	Adc49212 Rabbit hy
20	273	25.2	1662	8 ACC51023	Acc51023 Human bla

21	273	25.2	1662	8 ABX76415	Abx76415 Lung canc
22	273	25.2	1662	11 ADN39051	Adn39051 Cancer/an
23	273	25.2	4049	8 ABX76214	Abx76214 Lung canc
24	273	25.2	4049	12 ADN05115	Adn05115 Antipgori
25	273	25.2	4220	12 ADJ74905	Adj74905 Marker ge
26	250	23.1	1767	10 ADD93927	Add93927 Xenopus l
27	236.4	21.8	1665	2 AAV18822	Aav18822 Murine hy
28	236.4	21.8	1665	3 AAZ88201	Aaz88201 Mouse hya
29	236.4	21.8	1665	8 ABZ76736	Abz76736 Mouse hya
30	236.4	21.8	1665	10 AAD59444	Aad59444 Mouse hya
31	236.4	21.8	5919	12 ADJ75814	Adj75814 Marker ge
32	207.4	19.2	1229	12 ADJ74906	Adj74906 Marker ge
33	179.4	16.6	1653	6 ABL41013	Ab141013 Murine hy
34	179.4	16.6	1659	8 ABZ76735	Abz76735 Mouse hya
35	179.4	16.6	1659	10 AAD59443	Aad59443 Mouse hya
36	179.4	16.6	2947	2 AAV18821	Aav18821 Murine hy
37	179.4	16.6	2948	6 AAI98697	Aai98697 Mouse isc
38	179.4	16.6	4194	9 ADA49685	Ada49685 DNA encod
39	179.4	16.6	4194	10 ADC59298	Adc59298 Mouse HAS
40	165.8	15.3	2890	3 AAZ88200	Aaz88200 Mouse hya
41	153.6	14.2	1659	10 ADL13696	Adl13696 Osteoarth
42	153.6	14.2	3003	6 ABK84373	Abk84373 Human CDN
43	153.6	14.2	3003	9 ADA49683	Ada49683 DNA encod
44	153.6	14.2	3003	10 ADC59296	Adc59296 Human HAS
45	153.6	14.2	3003	10 ADD12718	Add12718 Human CDN

ALIGNMENTS

RESULT 1
ADL13691
ID ADL13691 standard; DNA; 1737 BP.
XX
AC ADL13691;

DT 06-MAY-2004 (first entry)

DE Osteoarthritis-associated polymorphic nucleotide #223.

XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.

OS Homo sapiens.

XX WO20003054166-A2.

XX 03-JUL-2003.

PF 19-DEC-2002; 2002WO-US041225.

XX 20-DEC-2001; 2001US-0342603P.

PR (INCY-) INCYTE GENOMICS INC.

XX Jones KA, Schafer A;

XX WPI; 2003-559141/52.

PT Determining susceptibility of an individual to joint space narrowing,
PT osteophyte development and/or joint pain comprises identifying whether
PT the individual has at least one polymorphism in a polynucleotide encoding
PT a protein.

XX Disclosure; SEQ ID NO 223; 297pp; English.

CC The invention relates to a method of determining susceptibility of an
CC individual to joint space narrowing and/or osteophyte development and/or
CC joint pain comprising identifying whether the individual has at least one
CC polymorphism in a polynucleotide encoding at least one of the protein
CC listed in the specification. The methods, composition and agent are
CC useful for modulating the susceptibility of an individual to joint space

CC narrowing and/or osteophyte development and/or joint pain that is
CC associated with a disease, preferably osteoarthritis. The cell line and
CC the non-human animal are useful for screening for an agent for diagnosing
CC an individual having susceptibility to joint space narrowing and/or
CC osteophyte development and/or joint pain. This sequence corresponds to
CC the polynucleotide encoding a protein listed in the specification. (Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC http://www.wipo.int/pub/published_pct_sequences/).

Sequence 1737 BP; 232 A; 571 C; 603 G; 331 T; 0 U; 0 Other;

Query Match 86.1%; Score 933; DB 10; Length 1737;

Best Local Similarity 90.6%; Pred. No. 2e-151;

Matches 1048; Conservative 0; Mismatches 35; Indels 74; Gaps 2;

Qy	1	ATGAGACAGCAGGAGCGGCCCAAGCCCACTCCTGAGACCCGCGCGCTGCTCGCGCCCTGGCC	60
Db	1	ATGAGACAGCAGGAGCGGCCCAAGCCCACTCCTGAGACCCGCGCGCTGCTCGCGCCCTGGCC	60
Qy	61	CGGAGGCTGCTGACCATCGCTTGGCCCTGTCTATCTGGGCTCATGACCTGGGCGCTTAC	120
Db	61	CGGAGGCTGCTGACCATCGCTTGGCCCTGTCTATCTGGGCTCATGACCTGGGCGCTTAC	120
Qy	121	GCGCCCGGGTGCCTGCTGGCTTCGATCGCTACGGCTCTCTGGCTTTCGGCGCTTACGGG	180
Db	121	GCGCCCGGGTGCCTGCTGGCTTCGATCGCTACGGCTCTCTGGCTTTCGGCGCTTACGGG	180
Qy	181	GCCTTCTTTTCAGCGCACCTTGGTGGCGCAGAGCTCTTTCGGGTACCTGGAGCACCGGGG	240
Db	181	GCCTTCTTTTCAGCGCACCTTGGTGGCGCAGAGCTCTTTCGGGTACCTGGAGCACCGGGG	240
Qy	241	GTGGCGGCGCGCGCGGGGCGCTGTGATGCAGCACCGCGCGCAGTGTGGCGCTGACC	300
Db	241	GTGGCGGCGCGCGCGGGGCGCTGTGATGCAGCACCGCGCGCAGTGTGGCGCTGACC	300
Qy	301	ATCTCCGCTTACAGAGGAGACCCCGCGTACTTGGCCAGTGTGGCGCTGACC	360
Db	301	ATCTCCGCTTACAGAGGAGACCCCGCGTACTTGGCCAGTGTGGCGCTGACC	360
Qy	361	CTGCTGTATCCCGCGCGCGCTGCGCTCTCATGTGTGGTGGATGGCAACCGCGCCGAG	420
Db	361	CTGCTGTATCCCGCGCGCGCTGCGCTCTCATGTGTGGTGGATGGCAACCGCGCCGAG	420
Qy	421	GACCTCTACATGGTTCGCGAGTCTTTCGTGACGAGGACCCCGCCACGTAC	480
Db	421	GACCTCTACATGGTTCGCGAGTCTTTCGTGACGAGGACCCCGCCACGTAC	480
Qy	481	GTGTGGGACGGCACTACCAACGCGCTGGGACCCGCGCGGGCGGGCGGGTGGCGCC	540
Db	481	GTGTGGGACGGCACTACCAACGCGCTGGGACCCGCGCGGGCGGGCGGGTGGCGCC	540
Qy	541	GGAGCTTATCGGAGGTGGAGCGAGGATCTGGGCGGTGGGCAAGTGGAGCGCTGGTG	600
Db	541	GGAGCTTATCGGAGGTGGAGCGAGGATCTGGGCGGTGGGCAAGTGGAGCGCTGGTG	600
Qy	601	AGGACTCGAGGTGCTGTGCTGGCGCAGCGCTGGGCGGCAAGCGAGGTCTATGTAC	660
Db	601	AGGACTCGAGGTGCTGTGCTGGCGCAGCGCTGGGCGGCAAGCGAGGTCTATGTAC	660
Qy	661	ACAGCTTTCAGGCGCTCGGAGATTCGCTGGATACGTGCAAGTCTGTGATCTCGGACACA	720
Db	661	ACAGCTTTCAGGCGCTCGGAGATTCGCTGGATACGTGCAAGTCTGTGATCTCGGACACA	720
Qy	721	AGGTTGGACCCCATGGCACTGCTGGAGTCTGTGGGGTACTGGACGAGGACCCCGGGTAC	780
Db	721	AGGTTGGACCCCATGGCACTGCTGGAGTCTGTGGGGTACTGGACGAGGACCCCGGGTAC	780
Qy	781	GGGCGTGTGGTGGGAGCGTGGGATCTCTTAAACCTCTGGACTCTCTGGGTGAGCTTCTTA	840
Db	781	GGGCGTGTGGTGGGAGCGTGGGATCTCTTAAACCTCTGGACTCTCTGGGTGAGCTTCTTA	840
Qy	841	AGCAGCTTGCATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTTCAGAGCTACTTCCAC	900

CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences/.

XX SQ Sequence 2088 BP; 302 A; 652 C; 711 G; 423 T; 0 U; 0 Other;

Query Match 86.1%; Score 933; DB 10; Length 2088;
Best Local Similarity 90.6%; Pred. No. 2e-151;
Matches 1048; Conservative 0; Mismatches 35; Indels 74; Gaps 2;

QY 1 ATGACAGCAGGAGCGCCCAAGCCACTCTCTGACGCGCGCGCTGCTCGGCGCTGGCC 60
DB 36 ATGACAGCAGGAGCGCGCCCAAGCCACTCTCTGACGCGCGCGCTGCTCGGCGCTGGCC 95
QY 61 CGGAGGGTGTACACCATCGCTTCGCGCTGCTCATCTCGGCGCTCATGACCTGGCGCTTAC 120
DB 96 CGGAGGGTGTACACCATCGCTTCGCGCTGCTCATCTCGGCGCTCATGACCTGGCGCTTAC 155
QY 121 GCGCGGGGTGCGCGCTGCTCCGATCGCTACCGGCTCTCGGCGCTTGGCGCTTCTACGGG 180
DB 156 GCGCGGGGTGCGCGCTGCTCCGATCGCTACCGGCTCTCGGCGCTTGGCGCTTCTACGGG 215
QY 181 GCGCTTCTTACGCGCACTGCTGGCGGAGAGCCCTTTCGCTACCTGGAGCACCGCGGG 240
DB 216 GCGCTTCTTACGCGCACTGCTGGCGGAGAGCCCTTTCGCTACCTGGAGCACCGCGGG 275
QY 241 GTGGCGGGCGCGCGCGCGCGCTGATGAGCCACCGCGCGAGTGTGCGCTGACC 300
DB 276 GTGGCGGGCGCGCGCGCGCGCTGATGAGCCACCGCGCGAGTGTGCGCTGACC 335
QY 301 ATCTCCGCTTACGAGAGAACCCCGCTACCTGCGCCAGTGCTCGGCGTCCGCGCGCC 360
DB 336 ATCTCCGCTTACGAGAGAACCCCGCTACCTGCGCCAGTGCTCGGCGTCCGCGCGCC 395
QY 361 CTGCTGTACCGCGCGCGCGCGCTGCTGCTCATGCTGGTGGTGGTGGTGGTGGTGGTGG 420
DB 396 CTGCTGTACCGCGCGCGCGCGCTGCTGCTCATGCTGGTGGTGGTGGTGGTGGTGGTGG 455
QY 421 GACCTCTACATGCTGCAATGCTTCGCGAGGCTTTCGCTGAGAGAGCCCGCCACGCTAC 480
DB 456 GACCTCTACATGCTGCAATGCTTCGCGAGGCTTTCGCTGAGAGAGCCCGCCACGCTAC 515
QY 481 GTGTGGGACGGCAATCAACAGCCCTGGGAAACCCCGCGCGCGCGCGCGCGCGCGCGCG 540
DB 516 GTGTGGGACGGCAATCAACAGCCCTGGGAAACCCCGCGCGCGCGCGCGCGCGCGCGCG 575
QY 541 GGAGCTTACGAGAGGTGGAGCGAGGATCTCTGGGCGGCTGGCAGTGGAGCGCGTGGTG 600
DB 576 GGAGCTTACGAGAGGTGGAGCGAGGATCTCTGGGCGGCTGGCAGTGGAGCGCGTGGTG 635
QY 601 AGGACTCGCAGGTGCTGTGCTGCGCAGCGCTGGGCGGCGCAAGCGGAGGTCACTGTAC 660
DB 636 AGGACTCGCAGGTGCTGTGCTGCGCAGCGCTGGGCGGCGCAAGCGGAGGTCACTGTAC 695
QY 661 ACAGCCTTCAAGCGCTCGGAGATTCGGTGGTACCTGCTGAGTCTGCTGGTTCCTA 720
DB 696 ACAGCCTTCAAGCGCTCGGAGATTCGGTGGTACCTGCTGAGTCTGCTGGTTCCTA 755
QY 721 AGTTGGACCCCATGGCACTGCTGGAGCTGCTGGGCGGCTGGAGCGAGGACCCCGGTA 780
DB 756 AGTTGGACCCCATGGCACTGCTGGAGCTGCTGGGCGGCTGGAGCGAGGACCCCGGTA 815
QY 781 GGGGCTGTGGTGGGAGCTGGGATCTTAACTCTGAGCTCTGGTTCCTA 840
DB 816 GGGGCTGTGGTGGGAGCTGGGATCTTAACTCTGAGCTCTGGTTCCTA 875
QY 841 AGCAGCTTCGATCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGCTTCTCCAC 900
DB 876 AGCAGCTTCGATCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGAGCTTCTCCAC 935
QY 901 TGTGTATCTGATCAGCGGCTCTAG----- 928
DB 936 TGTGTATCTGATCAGCGGCTCTAGGCTTATAGGAATAACCTCTTGCAGCAGTTT 995

QY 929 -----AATCTGCCAGGCCCCCAGGAGCAGCGATGATGCC 965
DB 996 CTTGAGGCTGTGTACCAACAGAGTTCTCTGGTACCCACTGTACTTTTGGGATGACCGG 1055
QY 966 CTCATTCTCTCGCCC-----CGTCAGGTACACCTCCAGGTCCCGC 1006
DB 1056 CACCTTCAACACCGCATGCTCAGCATGGTTATGTACCAAGTACACCTCCAGGTCCCGC 1115
QY 1007 TGCTACTCAGAGAGCGCCCTGCTCTCTCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 1066
DB 1116 TGCTACTCAGAGAGCGCCCTGCTCTCTCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 1175
QY 1067 AAGTCGTACTTCCGTGA 1083
DB 1176 AAGTCGTACTTCCGTGA 1192

RESULT 3

AAT99541
ID AAT99541 standard; cDNA; 2116 BP.

XX AC AAT99541;

XX DT 21-MAY-1998 (first entry)

XX XX Human hyaluronan synthase cDNA clone 30C.

DE DE Hyaluronan synthase; HAS gene; human; hyaluronic acid; cell adhesion;

KW KW wound healing; vulnery; tissue repair; scar; keloid; therapy; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 36..1172

FT polyA_signal /*tag= a

FT /*tag= b

XX PN WO9740174-A1.

XX PD 30-OCT-1997.

XX PF 17-APR-1997; 97WO-US006350.

XX PR 22-APR-1996; 96US-00635552.

XX PA (LEUK-) LEUKOSITE INC.

XX PI Brieskin MJ;

XX DR WPI; 1997-549359/50.

XX DR P-PSDB; AAW26765.

XX XX Human hyaluronan synthase - useful for recombinant production of

PT hyaluronic acid for wound healing, tissue repair and reducing

PT hypertrophic scar and keloid formation.

XX PS Claim 3; Page 36-38; 58pp; English.

CC CC cDNA clone 30C includes a coding region for hyaluronan synthase (HAS)

CC (see AAW26765), an enzyme involved in the synthesis of hyaluronan

CC (hyaluronic acid) and which has the ability to confer cell adhesion by

CC the lymphocyte receptor CD44. Clone 30C was isolated using an expression

CC cloning system developed to isolate cDNA clones that encode proteins that

CC confer adhesion of the murine T cell lymphoma TK1. A human mesenteric

CC lymph node expression library was constructed that, upon transfection

CC into CHO/p cells, yielded clone 30C that mediated rosetting of TK1 cells

CC to some of the transfectants. The isolated clone can be utilised in a

CC claimed method for producing HAS in host cells. Such host cells are used

CC in a claimed method for the production of hyaluronan. Hyaluronan is

CC useful for wound healing and tissue repair, and can reduce or prevent

CC hypertrophic scars and keloid formation. It is also used in eye surgery

CC as a replacement for vitreous fluid

SQ Sequence 2116 BP; 330 A; 651 C; 712 G; 423 T; 0 U; 0 Other;

Query Match 86.1%; Score 933; DB 2; Length 2116;
Best Local Similarity 90.6%; Pred. No. 2e-151;
Matches 1048; Conservative 0; Mismatches 35; Indels 74; Gaps 2;

QY 1 ATGACAGCAGGAGCGCCAGCCCACTCTGTCAGCGCCCGCTGCTCCGGCTGGCC 60
DB 36 ATGACAGCAGGAGCGCCAGCCCACTCTGTCAGCGCCCGCTGCTCCGGCTGGCC 95
QY 61 CGGAGGGTGTCTACCATCGCTTCGCCCTTCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
DB 96 CGGAGGGTGTCTACCATCGCTTCGCCCTTCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 155
QY 121 GCCCGCGGGTGCCTGCTGCTCCGATCGCTACAGCCCTCTTGCGCTTTCGGCTCTTACGGG 180
DB 156 GCCCGCGGGTGCCTGCTGCTCCGATCGCTACAGCCCTCTTGCGCTTTCGGCTCTTACGGG 215
QY 181 GCCTTCCTTCAGCGCACTGTGTGGCGAGAGCTTTCGGGTACTGTGAGCAGCGCGG 240
DB 216 GCCTTCCTTCAGCGCACTGTGTGGCGAGAGCTTTCGGGTACTGTGAGCAGCGCGG 275
QY 241 GTGGCGCGCGCGCGCGCGCGCTGCTGATGAGCCAGCGCGCGCTGCTGCGCTGACC 300
DB 276 GTGGCGCGCGCGCGCGCGCGCTGCTGATGAGCCAGCGCGCGCTGCTGCGCTGACC 335
QY 301 ATCTCCGCTTACAGAGGAGACCCCGCTACCTGCGCCAGTGTGCGCTTCGCCCGCGCC 360
DB 336 ATCTCCGCTTACAGAGGAGACCCCGCTACCTGCGCCAGTGTGCGCTTCGCCCGCGCC 395
QY 361 CTGCTGTACCGCGCGCGCGCTGCTGCTCTCTCATGCTGTGGATGGCAACCGGCGCGAG 420
DB 396 CTGCTGTACCGCGCGCGCGCTGCTGCTCTCTCATGCTGTGGATGGCAACCGGCGCGAG 455
QY 421 CACCTCTACATGCTGCACATGCTTCGCGAGGCTCTTCGCTGACGAGACCCCGCCACAGTAC 480
DB 456 GACCTCTACATGCTGCACATGCTTCGCGAGGCTCTTCGCTGACGAGACCCCGCCACAGTAC 515
QY 481 GTGTGGGACGGCAACTACACAGCCCTTGGGAAACCCCGCGCGCGCGCTGGGCGCC 540
DB 516 GTGTGGGACGGCAACTACACAGCCCTTGGGAAACCCCGCGCGCGCGCTGGGCGCC 575
QY 541 GGAGCCTATCGGAGGTGGAGCGGAGGATCTTGGGCGGCTGGCAGTGGAGGCGCTGGTG 600
DB 576 GGAGCCTATCGGAGGTGGAGCGGAGGATCTTGGGCGGCTGGCAGTGGAGGCGCTGGTG 635
QY 601 AGGACTCGCAGTGTGCTGTCGCGCAGCGCTGGGCGCGCAAGCGGAGGTATGTATAC 660
DB 636 AGGACTCGCAGTGTGCTGTCGCGCAGCGCTGGGCGCGCAAGCGGAGGTATGTATAC 695
QY 661 ACAGCCTTCAAGCGGCTCGGAGATTCGCTGAGTACGTCAGAGTCTGTGACTCGGACACA 720
DB 696 ACAGCCTTCAAGCGGCTCGGAGATTCGCTGAGTACGTCAGAGTCTGTGACTCGGACACA 755
QY 721 AGGTGGACCCCATGCGACTGCTGAGCTCTGCGGGTACTGTGAGGAGGAGACCCCGGGTA 780
DB 756 AGGTGGACCCCATGCGACTGCTGAGCTCTGCGGGTACTGTGAGGAGGAGACCCCGGGTA 815
QY 781 GGGGCTGTGTGGGAGCTGCGGATCTTAACCTCTGGACTCTCTGGGTGAGTCTTCTTA 840
DB 816 GGGGCTGTGTGGGAGCTGCGGATCTTAACCTCTGGACTCTCTGGGTGAGTCTTCTTA 875
QY 841 AGCAGCTTCGATACTTGGGTAGCTTCAATGTGAGCGGGCTTGTGAGAGTACTTTCAC 900
DB 876 AGCAGCTTCGATACTTGGGTAGCTTCAATGTGAGCGGGCTTGTGAGAGTACTTTCAC 935
QY 901 TGTGTATCTGCATCAGCGGTCTCTAG----- 928
DB 936 TGTGTATCTGCATCAGCGGTCTCTAGGCCTATATAGGAATAACCTTTGACAGAGTTT 995
QY 929 -----AATCTGCCAGGGCCCGGAGGAGCAGCGATGATGCC 965

DB 996 CTTGAGGCTGTGTACAAACAGAGTTCTCTGGTACCCACTGTACTTTTGGGGATGACCGG 1055
QY 966 CTGATTTCTGCCCC-----CGTGAGGTACACCTCCAGGTCCCGC 1006
DB 1056 CACCTCAACCAACCGCATGCTCAGCATGGGTTATGCTACCAAGTACACCTCCAGGTCCCGC 1115
QY 1007 TGCTACTCAGAGAGCGCTCTGCTCTCTGCGGTGGCTGAGCCAGACACACGCTGGTCC 1066
DB 1116 TGCTACTCAGAGAGCGCTCTGCTCTCTGCGGTGGCTGAGCCAGACACACGCTGGTCC 1175
QY 1067 AAGTCGTACTTCCGTGA 1083
DB 1176 AAGTCGTACTTCCGTGA 1192

RESULT 4

ADL13692
ID ADL13692 standard; DNA; 1737 BP.

XX ADL13692;

AC ADL13692;

DT 06-MAY-2004 (first entry)

XX Osteoarthritis-associated polymorphic nucleotide #224.

XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;

XX joint space narrowing; osteophyte development; joint pain;

XX osteoarthritis; SNP; single nucleotide polymorphism.

XX Homo sapiens.

XX WO2003054166-A2.

XX 03-JUL-2003.

XX 19-DEC-2002; 2002WO-US041225.

XX 20-DEC-2001; 2001US-0342603P.

XX (INCY-) INCYTE GENOMICS INC.

XX Jones KA, Schafer A;

XX WPI; 2003-559141/52.

XX Determining susceptibility of an individual to joint space narrowing,

XX osteophyte development and/or joint pain comprises identifying whether

XX the individual has at least one polymorphism in a polynucleotide encoding

XX a protein.

XX Disclosure; SEQ ID NO 224; 297pp; English.

XX The invention relates to a method of determining susceptibility of an
XX individual to joint space narrowing and/or osteophyte development and/or
XX joint pain comprising identifying whether the individual has at least one
XX polymorphism in a polynucleotide encoding at least one of the protein
XX listed in the specification. The methods, composition and agent are
XX useful for modulating the susceptibility of an individual to joint space
XX narrowing and/or osteophyte development and/or joint pain that is
XX associated with a disease, preferably osteoarthritis. The cell line and
XX the non-human animal are useful for screening for an agent for diagnosing
XX an individual having susceptibility to joint space narrowing and/or
XX osteophyte development and/or joint pain. This sequence corresponds to
XX the polynucleotide encoding a protein listed in the specification. (Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences).

XX Sequence 1737 BP; 232 A; 569 C; 604 G; 332 T; 0 U; 0 Other;

Query Match 86.0%; Score 931.4; DB 10; Length 1737;
Best Local Similarity 90.5%; Pred. No. 3.8e-151;

Qy 1 ATGAGACAGCAGGACGCGCCCCAAGCCCACTCCTGCAGCCCGCCGCTGCTCCGGCCCTGGCC 60

Db 36 ATGAGACAGCAGGACGCGCCCCAAGCCCACTCCTGCAGCCCGCCGCTGCTCCGGCCCTGGCC 95

QY 61 CGAGGGGTGCTGACCATGCGCTTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCGCTTAC 120
DB |||||||
QY 96 CGAGGGGTGCTGACCATGCGCTTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCGCTTAC 155
DB |||||||
QY 121 GCCCGGGGGTGCCTGCTGCTCCGATCGCTACGGCTCTCTGGGCTCTGGGCTCTTACGGG 180
DB |||||||
QY 156 GCCCGGGGGTGCCTGCTGCTCCGATCGCTACGGCTCTCTGGGCTCTGGGCTCTTACGGG 215
DB |||||||
QY 181 GCCTTCCTTTTACAGCGACCTGCTGGGCGAGAGCTCTTTCGCTACTGAGAGCACGGGCGG 240
DB |||||||
QY 216 GCCTTCCTTTTACAGCGACCTGCTGGGCGAGAGCTCTTTCGCTACTGAGAGCACGGGCGG 275
DB |||||||
QY 241 GTGGCGGGGGCGCGCGGGGCGCTGATGACGACCGCGCGAGTGTGCGCTGACC 300
DB |||||||
QY 276 GTGGCGGGGGCGCGCGGGGCGCTGATGACGACCGCGCGAGTGTGCGCTGACC 335
DB |||||||
QY 301 ATCTCCGGCTTACAGAGGACCCCGCTACCTGCGCCAGTGTCTGGGCTCTGGGCGCGGCC 360
DB |||||||
QY 336 ATCTCCGGCTTACAGAGGACCCCGCTACCTGCGCCAGTGTCTGGGCTCTGGGCGCGGCC 395
DB |||||||
QY 361 CTGCTGTATACCGCGCGCGGGTCTGGGCTCTCATGCTGTGGATGGCAACCGCGCGGAG 420
DB |||||||
QY 396 CTGCTGTATACCGCGCGCGGGTCTGGGCTCTCATGCTGTGGATGGCAACCGCGCGGAG 455
DB |||||||
QY 421 GACCTCTATAGTGTGACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCCACGTAC 480
DB |||||||
QY 456 GACCTCTATAGTGTGACATGTTTCGCGAGGTCTTTCGCTGACGAGGACCCCGCCACGTAC 515
DB |||||||
QY 481 GTGTGGGACGGCACTACACAGCCCTTGGGAACCCCGGGCGGGCGCGGTGGGCGGCC 540
DB |||||||
QY 516 GTGTGGGACGGCACTACACAGCCCTTGGGAACCCCGGGCGGGCGCGGTGGGCGGCC 575
DB |||||||
QY 541 GGAGCCTATCGGAGGCTGAGCGAGAGTCTTGGGCGGGCTGGCAGTGGAGCGCGTGTGTG 600
DB |||||||
QY 576 GGAGCCTATCGGAGGCTGAGCGAGAGTCTTGGGCGGGCTGGCAGTGGAGCGCGTGTGTG 635
DB |||||||
QY 601 AGGACTCCGAGTGTGCTGCGCGAGCGCTGGGCGGGCGCAAGCGAGGTCTATGTAC 660
DB |||||||
QY 636 AGGACTCCGAGTGTGCTGCGCGAGCGCTGGGCGGGCGCAAGCGAGGTCTATGTAC 695
DB |||||||
QY 661 ACAGCCTTCAAGCGCTCGGAGATTTCGGTGGACTACGTGCGAGGTCTGTGACTCGGACACA 720
DB |||||||
QY 696 ACAGCCTTCAAGCGCTCGGAGATTTCGGTGGACTACGTGCGAGGTCTGTGACTCGGACACA 755
DB |||||||
QY 721 AGTTGGAGCCCATGCGACTGCTGAGCTCTGCGGGTACTGGAGAGAGCCCGGGTA 780
DB |||||||
QY 756 AGTTGGAGCCCATGCGACTGCTGAGCTCTGCGGGTACTGGAGAGAGCCCGGGTA 815
DB |||||||
QY 781 GGGGCTGTGTGGGAGCTGCGGATCTTAACCTCTGGAATCTGGAATCTGGAATCTGGAATCTGGA 840
DB |||||||
QY 816 GGGGCTGTGTGGGAGCTGCGGATCTTAACCTCTGGAATCTGGAATCTGGAATCTGGAATCTGGA 875
DB |||||||
QY 841 AGCAGCTGCGATCTGCGGTAGCTTCAATGTGAGCGGGTGTGTGAGAGTACTTCCAC 900
DB |||||||
QY 876 AGCAGCTGCGATCTGCGGTAGCTTCAATGTGAGCGGGTGTGTGAGAGTACTTCCAC 935
DB |||||||
QY 901 TGTGTATCTGATCAGCGGTCTCTAG----- 928
DB |||||||
QY 936 TGTGTATCTGATCAGCGGTCTCTAGGCCCTATATAGGAATAACCTCTTGCAGCAGTTT 995
DB |||||||
QY 929 -----AATCCTGCCAGGCCCGGAGGACGCGATGATGCC 965
DB |||||||
QY 996 CTTGAGGCTGGTACAACAGAGTTCTTGGGTACCCACTGTATCTTTTGGGGATGACCGG 1055
DB |||||||
QY 966 CTCATTCCTCGCCCC-----CGTGCAGGTACACCTCCAGGTCCCGC 1006
DB |||||||
QY 1056 CAGCTCAACCAACCGCATGCTCAGCATGGTTATGTACCAAGTACACCTCCAGGTCCCGC 1115
DB |||||||
QY 1007 TGCTACTCAGAGAGCCCTCTGCTCTTCTGCGGTGGCTGAGCGAGACAGACGCTGTGTC 1066
DB |||||||
QY 1116 TGCTACTCAGAGAGCCCTCTGCTCTTCTGCGGTGGCTGAGCCAGACAGACGCTGTGTC 1175
DB |||||||
QY 1067 AAGTCGTACTTCCGTGA 1083

DB 1176 AAGTCGTACTTCCGTGA 1192
|||||
RESULT 6
AAT96713
ID AAT96713 standard; DNA; 2117 BP.
XX AAT96713;
AC AAT96713;
XX
DT 22-APR-1998 (first entry)
XX
XX Human hyaluronate synthetase coding sequence.
DE
XX Hyaluronate synthetase; human; hyaluronic acid; drug preparation;
KW cosmetic preparation; gene therapy; carcinogenesis; ss.
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT CDS 149..1780
FT /*tag= a
XX
XX WO9738113-A1.
PN
PD 16-OCT-1997.
XX
XX 31-MAR-1997; 97WO-JP001111.
PF
XX 05-APR-1996; 96JP-00084326.
PR
XX 30-APR-1996; 96JP-00109663.
XX
XX (SEBK) SEIKAGAKU CORP.
PA
XX Itano N, Kimata K;
PI
XX WPI; 1997-512726/47.
DR P-PSDB; AAW36503.
XX
XX DNA encoding human hyaluronate synthetase - for industrial scale
PT production of hyaluronic acid used in generating anti-carcinogenic drugs
PT or for cosmetics.
XX
XX Claim 3; Page 23-27; 35pp; Japanese.
XX
XX This sequence encodes a human hyaluronate synthetase, and is the coding
CC sequence of the invention. The encoded enzyme is useful for industrial
CC scale production of hyaluronic acid for use in the preparation of drugs
CC and cosmetics. The drugs can also be used in compositions for the
CC treatment of disorders involving the lowering of hyaluronic acid
CC production. The peptides may be used for the preparation of antibodies
CC recognising hyaluronate synthetase, e.g. for diagnostic purposes.
CC Antisense DNA or RNA corresponding to the DNA can be used in gene therapy
CC treatment of carcinogenesis.
XX
XX Sequence 2117 BP; 306 A; 673 C; 703 G; 435 T; 0 U; 0 Other;
SQ
Query Match 84.1%; Score 911; DB 2; Length 2117;
Best Local Similarity 89.7%; Pred. No. 1.2e-147;
Matches 1032; Conservative 0; Mismatches 45; Indels 74; Gaps 2;
QY 7 CAGCAGAGCGCGCCAGCCCACTCTCGAGCGCGCGCTGCTCGGCTGGCGCGGAGG 66
DB |||||||
QY 50 CAGCAGAGCGCGCCAGCCCACTCTCGAGCGCGCGCTGCTCGGCTGGCGCGGAGG 109
DB |||||||
QY 67 GTGCTGACCATCGGCTTTCGGCTCTCATCTGGGCTCATGACCTGGGCTAGCGCGCC 126
DB |||||||
QY 110 GTGCTGACCATCGGCTTTCGGCTCTCATCTGGGCTCATGACCTGGGCTAGCGCGCC 169
DB |||||||
QY 127 GGGGTGCGCTGGGCTTCGATCGCTACGGCTCTCTGGGCTTTCGGGCTTTC 186
DB |||||||
QY 170 GGGGTGCGCTGGGCTTCGATCGCTACGGCTCTCTGGGCTTTCGGGCTTTC 229
DB |||||||

Qy		187	CTTTCAGCGCACCTGTGGCGAAGAGCCTCTTTCGCGTAACCTGGAGACAACCGCGGGGTGGCG	246
Db		230	CTTTTCAGCGCACCTGTGTGGCCACAGAGACCTCTTTCGCGTAACCTGGAGACAACCGCGGGGTGGCG	289
Qy		247	GCGCGCGCGGGGGCGCTTGGAATGCAGCCACCGCGCACTGTGGCGCTGACCATCTCC	306
Db		290	GCGCGCGCGGGGGCGCTTGGAATGCAGCCACCGCGCACTGTGGCGCTGACCATCTCC	349
Qy		307	GCCTAC CAGGAGGACCCCGGCTACTTCGCGCAAGTGCCTGGCGTCCGCCCGCCCTGCTG	366
Db		350	GCCTAC CAGGAGGACCCCGGCTACTTCGCGCAAGTGCCTGGCGTCCGCCCGCCCTGCTG	409
Qy		367	TACCCGCGCGGGCTGGCGGCTCTCATGGTGGTGGATGGCAAACCGCGCGGAGACCTC	426
Db		410	TACCCGCGCGGGCTGGCGGCTCTCATGGTGGTGGATGGCAAACCGCGCGGAGACCTC	469
Qy		427	TACATGCTGCACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCGCACTGACGTGTGG	486
Db		470	TACATCGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCGCACTGACGTGTGG	529
Qy		487	GACGGCAACTACACACAGCCCTGGGAAACCGCGCGCGCGGGCGGTGGGGCGCGGAGCC	546
Db		530	GACGGCAACTACACACAGCCCTGGGAAACCGCGCGCGCGGGCGGTGGGGCGCGGAGCC	589
Qy		547	TATCGGAGGTGGAGGCGGAGGATCCCTGGCGGCTGCAGTGGAGGGCGCTGGTAGGACT	606
Db		590	TATCGGAGGTGGAGGCGGAGGATCCCTGGCGGCTGCAGTGGAGGGCGCTGGTAGGACT	649
Qy		607	CGCAGGTGCGTGTGCGTGGGCGACGCGTGGGGGGCAAGCGCGAGGTCACTGTACACAGCC	666
Db		650	CGCAGGTGCGTGTGCGTGGGCGACGCGCTGGGGGGCAAGCGCGAGGTCACTGTACACAGCC	709
Qy		667	TTCAAGGCGCTCGGAGATTGCGTGGACTACGTGCAAGTCTGTGACTCGGACACAAAGTTG	726
Db		710	TTCAAGGCGCTCGGAGATTGCGTGGACTACGTGCAAGTCTGTGACTCGGACACAAAGTTG	769
Qy		727	GACCCCATGCATCTGCGAGCTCGTSCGGGTACTGGACGAGGACCCCGGGTAGGGGCT	786
Db		770	GACCCCATGCATCTGCGAGCTCGTSCGGGTACTGGACGAGGACCCCGGGTAGGGGCT	829
Qy		787	GTTGGTGGGACGTCGCGGATCCTTTAACCTCTCGAGCTCCTGGGTCACTTCTTAAGCAGC	846
Db		830	GTTGGTGGGATGTCGCGATCCTTTAACCTCTCGAGCTCCTGGGTCACTTCTTAAGCAGC	889
Qy		847	CTGGGATACTGGGTAGCCTTCAATGTGGAGGGGCTTGTGAGAGTACTTCCACTGTGTA	906
Db		890	CTGGGATACTGGGTAGCCTTCAATGTGGAGGGGCTTGTGAGAGTACTTCCACTGTGTA	949
Qy		907	TCCTGCATCAGCGTCCCTCTAG-----	928
Db		950	TCCTGCATCAGCGTCCCTTAGGCGCTATATAGGATAAACCTCTTGGACGAGTTCTTTGAG	1009
Qy		929	-----AATCCTGCCAGGCCCGCAGGAGCACGCGATGATGCCCTCAATT	971
Db		1010	GCCTGTGTAAAC CAGAGATTCTTGGGTACCACTGTACTTTTGGGATGACCGGACCTC	1069
Qy		972	CCTCGCCC-----CGTGCAGGTACACCTTCAGGTCCTCGGCTCCCGTGTCTAC	1012
Db		1070	ACCAACCGCATGCTCAGCATGGGTTATGCTACCAAGTACACCTTCAGGTCCTCGCTGTCTAC	1129
Qy		1013	TCAGAGACGCCCTCGTCTTCTCGCGTGGCTGAGCCAGCAGACACGCTGGTCCAAAGTCG	1072
Db		1130	TCAGAGACGCCCTCGTCTTCTCGCGTGGCTGAGCCAGCAGACACGCTGGTCCAAAGTCG	1189
Qy		1073	TACTTCCGTGA 1083	
Db		1190	TACTTCCGTGA 1200	

RESULT 7

RESOL 7
ADL73693/C

ADL13693/C
ID ADL13693 standard; DNA: 231222 BP.

XX

AC	ADL13693;	
XX	06-MAY-2004 (first entry)	
XX	Osteoarthritis-associated polymorphic nucleotide #225.	
XX	ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;	
KW	joint space narrowing; osteophyte development; joint pain;	
KW	osteoarthritis; SNP; single nucleotide polymorphism.	
XX	Homo sapiens.	
OS	WO2003054166-A2.	
XX	03-JUL-2003.	
XX	19-DEC-2002; 2002WO-US041225.	
PF	20-DEC-2001; 2001US-0342603P.	
XX	(INCY-) INCYTE GENOMICS INC.	
PA	Jones KA, Schafer A;	
PI	WPI; 2003-559141/52.	
DR	Determining susceptibility of an individual to joint space narrowing,	
XX	osteoarthrite development and/or joint pain comprises identifying whether	
PT	the individual has at least one polymorphism in a polynucleotide encoding	
PT	a protein.	
XX	Disclosure; SEQ ID NO 225; 297pp; English.	
XX	The invention relates to a method of determining susceptibility of an	
CC	individual to joint space narrowing and/or osteophyte development and/or	
CC	joint pain comprising identifying whether the individual has at least one	
CC	polymorphism in a polynucleotide encoding at least one of the protein	
CC	listed in the specification. The methods, composition and agent are	
CC	useful for modulating the susceptibility of an individual to joint space	
CC	narrowing and/or osteophyte development and/or joint pain that is	
CC	associated with a disease, preferably osteoarthritis. The cell line and	
CC	the non-human animal are useful for screening for an agent for diagnosing	
CC	an individual having susceptibility to joint space narrowing and/or	
CC	osteoarthrite development and/or joint pain. This sequence corresponds to	
CC	the polynucleotide encoding a protein listed in the specification. (Note:	
CC	The sequence data for this patent did not form part of the printed	
CC	specification but was obtained in electronic format directly from WIPO at	
CC	ftp.wipo.int/pub/published_pct_sequences).	
XX	Sequence 231222 BP; 64006 A; 52087 C; 53478 G; 61650 T; 0 U; 1 Other;	
QY	Query Match 64.4%; Score 597.2; DB 10; Length 231222;	
DB	Best Local Similarity 99.6%; Pred. No. 5.6e-111;	
DB	Matches 699; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
QY	7 CAGCAGACGCGCCCAAGCCACCTCTGCAGCCGCGCTGCTCGCGCTGCGCCGAGG 66	
DB	149122 CAGCAGACGCGCCCAAGCCACCTCTGCAGCTGCTCGCGCTGCGCCGAGG 149063	
QY	67 GTGCTGACCATCGCTTTCGCCCTCTCATCTTGGGCTCTATGACCTGGGCTTACGCCGCC 126	
DB	149062 GTGCTGACCATCGCTTTCGCCCTCTCATCTTGGGCTCTATGACCTGGGCTTACGCCGCC 149003	
QY	127 GGGGTGCGCTGGGCTCGATCGATCGGCTCTTGGGCTCTTGGGCTCTACGGGCGCTTC 186	
DB	149002 GGGGTGCGCTGGGCTCGATCGATCGGCTCTTGGGCTCTTGGGCTCTACGGGCGCTTC 148943	
QY	187 CTTTCAGCGACCTGGTGGGCGCAGAGCTCTTTCGCTACCTGGGACACCGGCGGTGGCG 246	
DB	148942 CTTTCAGCGACCTGGTGGGCGCAGAGCTCTTTCGCTACCTGGGACACCGGCGGTGGCG 148883	
QY	247 GCGGCGGCGGGGGCGCGCTGGATGACGCCACCGCGCGCAGTGTGGCGCTGACCATCTCC 306	

Db 148882 GCGGCGCGCGCGCGCTGGATGACGACCGCGCGAGTGTGGCGCTGACCATCTCC 148823
 QY 307 GCCTACAGAGGAGCCCGCGTACCTGGCGCCAGTGCCTGGCTCGCCCGCGCCCTGCTG 366
 Db 148822 GCCTACAGAGGAGCCCGCGTACCTGGCGCCAGTGCCTGGCTCGCCCGCGCCCTGCTG 148763
 QY 367 TACCGCGCGCGCGCGCTGGCTCTCATGTGTGGATGGCAACCGCGCGAGGACCTC 426
 Db 148762 TACCGCGCGCGCGCGCTGGCTCTCATGTGTGGATGGCAACCGCGCGAGGACCTC 148703
 QY 427 TACATGTGTCGATCTTCGCGAGGTCTTCGCTGACGAGGACCCCGCCACATGATGTGG 486
 Db 148702 TACATGTGTCGATCTTCGCGAGGTCTTCGCTGACGAGGACCCCGCCACATGATGTGG 148643
 QY 487 GACGCAACTACCAACGCTTGGAAACCGCGCGCGCGCGGTGGCGCGCGAGCC 546
 Db 148642 GACGCAACTACCAACGCTTGGAAACCGCGCGCGCGCGGTGGCGCGCGAGCC 148583
 QY 547 TATCGGAGGTGGAGCGGAGGATCCTGGCGGCTGGCAGTGGAGGCGCTGTGAGGACT 606
 Db 148582 TATCGGAGGTGGAGCGGAGGATCCTGGCGGCTGGCAGTGGAGGCGCTGTGAGGACT 148523
 QY 607 GCGAGTGCCTGTGCTGGCGCAGCGCTGGCGCGCGCGCGAGGTGATGTACACAGCC 666
 Db 148522 GCGAGTGCCTGTGCTGGCGCAGCGCTGGCGCGCGCGCGAGGTGATGTACACAGCC 148463
 QY 667 TTCAGGCGCTCGGAGATTCGGTGAATACGTGCAAGTCTGT 708
 Db 148462 TTCAGGCGCTCGGAGATTCGGTGAATACGTGCAAGTCTGT 148421

RESULT 8

ABZ76734
 ID ABZ76734 standard; cDNA; 1752 BP.
 XX
 AC ABZ76734;
 XX
 DT 01-APR-2000 (first entry)
 XX
 DE Mouse hyaluronan synthase 1 encoding cDNA SEQ ID NO:1.
 KW Mouse; hyaluronan synthase; HAS; HAS1; enzyme; ophthalmological;
 KW osteopathic; antiarthritic; gene therapy; angiogenesis inhibitor;
 KW hyaluronic acid; dihydrazide; bioconjugate; dry eye syndrome; eye;
 KW osteoarthritis; gene; ss.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1752
 FT /*tag= a
 FT /product= "hyaluronan synthase 1 (HAS1)"
 XX
 FN WO2003006068-A1.
 XX
 PD 23-JAN-2003.
 XX
 PF 10-JUL-2001; 2001WO-US021785.
 XX
 PR 10-JUL-2001; 2001WO-US021785.
 XX
 PA (CLEA-) CLEAR SOLUTIONS BIOTECH INC.
 XX
 PI Dehazya P, Chen W;
 XX
 DR WPI; 2003-221664/21.
 DR P-PSDB; ABP96028.
 XX
 PT Novel dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate
 PT for treating dry eye syndrome, has derivatized hyaluronic acid
 PT crosslinked to nucleic acid encoding protein with hyaluronan synthase
 PT activity.
 XX

PS Claim 19; Page 56; 62pp; English.
 XX
 CC The present invention describes a dihydrazide derivatised hyaluronic acid
 CC (HA)/nucleic acid (NA) bioconjugate (I), with derivatised HA linked to NA
 CC encoding protein with hyaluronan synthase (HAS) activity, where NA has
 CC sequence with 79-85% sequence identity (SI) to a sequence of 1752, 1659
 CC or 1665 nucleotides (see ABZ76734 to ABZ76736) or a sequence encoding
 CC proteins with 66-90% SI to a sequence of 583, 552 or 554 amino acids (see
 CC ABP96028 to ABP96030). (I) has ophthalmological, osteopathic and
 CC antiarthritic activities, and can be used in gene therapy and as an
 CC inhibitor of angiogenesis, and as an inducer of expression of (HA) in
 CC human corneal epithelial cell. (I) is useful for transfecting a cell of
 CC an eye with nucleic acid, where the nucleic acid comprises (ABZ76734 to
 CC ABZ76736), by contacting the cell with (I) comprising the nucleic acid.
 CC (I) is useful for treating dry eye syndrome in an individual. (I) is
 CC useful in gene therapy applications for the treatment of a variety of
 CC medical conditions including dry eye syndrome or other medical conditions
 CC where an increase in the production of (HA) in the eye would be
 CC therapeutic (e.g., osteoarthritis of the articular joints). (I) is also
 CC useful for inhibiting angiogenesis for the treatment of macular
 CC degeneration or genes related to lipid biosynthesis that helps to restore
 CC the lipid component of the tear film, and as reagents for in vitro
 CC transformation of any cell, preferably a eukaryotic cell, more preferably
 CC a human eye cell. The present sequence encodes mouse HAS1 which is used
 CC in the exemplification of the present invention
 XX
 SQ Sequence 1752 BP; 308 A; 505 C; 548 G; 391 T; 0 U; 0 Other;
 Query Match 62.6%; Score 677.8; DB 8; Length 1752;
 Best Local Similarity 84.5%; Pred. No. 1.4e-107;
 Matches 792; Conservative 0; Mismatches 127; Indels 18; Gaps 2;
 QY 10 CAGGACGCGCCAGCCCACTCTCCAGACCGCGCGCTGCTCCGCGCTGGCCCGAGGGTG 69
 Db 7 CAGGACATGCCAAAGCCCTCAGAGGCGAGCGGTTCCTGCTCTGGCCTGGCCAGCGAGCA 66
 QY 70 CTGACCATCGCTTCGCGCTGCTCATCTGCGGCTCATGACCTGGGCTACGCGCGCGGG 129
 Db 67 CTCAGATCATCTTTGCGCTGCTCATCTGCGGCTCATGACCTGGGCTACGCGCGAGGC 126
 QY 130 GTGCGCTGGCTCCGATCGCTACGGCTCTCTGCGCTTCGCGCTTACGCGGCTTCTCTT 189
 Db 127 GTTCCTCTGGCTTCAGATCGCTATGAGACTCTGCGCTTTGGCTCTATATGGGCTTCTCT 186
 QY 190 TCAGCGACCTGTGGCGCAGAGCTCTTCGCTACCTGAGACACCGCGGCTGGCGCG 249
 Db 187 AGCGCACACCTAGTGGCAGAGGCTCTTCGCTTACCTGGAGCACCGAAGGCTGGCAGCG 246
 QY 250 GCGGCGG-----GGGCGCGCTGATGCGACCGCGCGCGAGTGTGGCG 294
 Db 247 GCTGCGGCGGCTCTTTGGGAGAGGGGCCCCCTGGATGCGGCCACTGACGAGCGTGGCA 306
 QY 295 CTGACCATCTCCGCTTACAGAGGAGACCCCGCGTACCTGCGGCCAGTGCTGGCGTCCGCC 354
 Db 307 CTCACCATCTCAGCTTACCAAGAGGATCCCGCTTACCTGCGCCAGTGCTTGACCTCCGCG 366
 QY 355 CGGCGCTCTGTACCGCGCGCGCTGGCGCTCTCATGTGTGGTGGATGGCAACGCG 414
 Db 367 CGGCGCTTGTGTACCGCGCACAGAGGTTCAGCGCTGCTCATGTGTGGTGGAGCGGCAACGCG 426
 QY 415 GCGGAGGACCTTACATGCTGACATGCTTCGCGAGGTCTTCGCTGACGAGGACCCCGCG 474
 Db 427 GCTGAGGATCTGTACATGTTGGATGTTCCGAGAGTCTTCGCGGATGAGGACCCCGCG 486
 QY 475 ACGTACGCTGTGGGACGGCAACTACACAGCCCTTGGGAACCCCGCGCGG---CGGCGCG 531
 Db 487 ACTTATGTGTGGATGGCAACTACCATCAGCCCTTGGGAACCCAGCGAGGCTTACGGCGCT 546
 QY 532 GTGGCGCGGAGCGCTTATCGGAGGTGGAGCGGAGGATCTTGGCGCGCTGGCAGTGGAG 591
 Db 547 GTGCGTGAAGGTGCTTACCGGAGGTGGAGCGGAGGACCCCGGCGGCTTGGCGGTGGAG 606
 QY 592 GCGCTGTGAGGACTCGCAGGTGCTGTGCTGCGCAGCGCTGGGGCGGCAAGCGCGAG 651

Db 607 GCGCTGGTGAGAACACGACGAGTGCCTGCTGGCTCAGCGTTGGGGCGCAACGTCGAG 666
QY 652 GTCATGTATACACAGCCTTCAAGGCGCTCGAGAGATTGGTGGACTACGTGCAGGTCTGTGAC 711
Db 667 GTCATGTATACACAGCCTTCAAGGCGCTCGAGAGATTGGTGGACTACGTGCAGGTCTGTGAC 726
QY 712 TCGGACACAAGGTTGGACCCCATGGCACTGTGGAGCTCGTGGGTACTGGACGAGGAC 771
Db 727 TCAGACACAAGACTAGACCCCATGGCACTGTGGAGCTTGTGGAGTGTGGATGAAGAC 786
QY 772 CCGCGGTAGGGGCTGTGGTGGGACGTCGGATCCTTAACCTCTGGACTCCTGGGTC 831
Db 787 CCGCGGTAGGGGCTGTGGAGGGGATGTGAGGATCCTTAACCTCTGGACTCCTGGGTC 846
QY 832 AGCTTCTTAAGCAGCCTGGATGCTGGGTAGCTTCAATGTGAGCGGGCTTGTGAGAGC 891
Db 847 AGCTTCTTAAGCAGCCTTTCGATACGAGGTAGCTTCAATGTGAGAGAGCTTGTGAGAGC 906
QY 892 TACTTCCACTGTGTATCCTGTCATCAGCGGTCCTCTAG 928
Db 907 TACTTCCACTGTGTCTCCTGTCATCAGTGTCTCTGG 943

RESULT 9

AAD59442
ID AAD59442 standard; DNA; 1752 BP.
AC AAD59442;
XX
DT 18-DEC-2003 (first entry)
XX
DE Mouse hyaluronan synthase (HAS) 1 DNA.
XX
KW Mouse; hyaluronic acid; HA; hyaluronan synthase; HAS; dry eye syndrome;
KW eye cell; osteoarthritis; gene therapy; gene; db.
XX
OS Mus sp.
FH Key Location/Qualifiers
FT CDS 1..1752
FT FT /*tag= a
FT FT /product= "Mouse hyaluronan synthase 1"
XX
XX US2003087850-A1.
XX
XX 08-MAY-2003.
XX
XX 10-JUL-2001; 2001US-00902939.
XX
XX 10-JUL-2001; 2001US-00902939.
XX
XX (DEHA/) DEHAZYA P.
XX (CHEN/) CHEN W.
XX
XX Dehazya P, Chen W;
XX
XX WPI: 2003-755151/71.
XX P-PSDB; AAE39152.
XX
XX Dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate
XX comprising derivatized hyaluronic acid cross linked to nucleic acid
XX encoding hyaluronan synthase useful for treating dry eye syndrome.
XX
XX Claim 1; Page 12-13; 31pp; English.
XX
XX The invention relates to dihydrazide derivatised hyaluronic acid (HA) /
XX nucleic acid bioconjugate comprising derivatised HA cross linked to
XX nucleic acid encoding hyaluronan synthase (HAS). The invention is useful
XX for transfected an eye cell of an individual. It is useful for treating
XX dry eye syndrome and osteoarthritis of the particular joints. The
XX invention is also useful in gene therapy. The present sequence is mouse
XX HAS1 DNA

XX SQ Sequence 1752 BP; 308 A; 505 C; 548 G; 391 T; 0 U; 0 Other;
Query Match 62.6%; Score 677.8; DB 10; Length 1752;
Best Local Similarity 84.5%; Pred. No. 1.4e-107; Indels 18; Gaps 2;
Matches 792; Conservative 0; Mismatches 127;
QY 10 CAGGACGGCCCAAGCCACCTCTCGACGCCCGCTGCTCGCGGCTGCGCCGAGGGTG 69
Db 7 CAGGACATGTCACCAAGCCCTCAGAGCAGCGGCTGCTGCTCTGGCCTGGCCAGGCGAGCA 66
QY 70 CTGACCATCGCCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTAGCCCGCGGG 129
Db 67 CTCAGATCATCTTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTCATGCGCAGGC 126
QY 130 GTGCGCTGGGCTCGATCGGCTACGGCTCTGGGCTTTCGGCTCTAGGGGCTTCTCTT 189
Db 127 GTTCTCTGGGCTTCAGATCGGCTATGGAATCTCTGGCTTTGGCTCTATGGGGCATTCCTC 186
QY 190 TCAGCGCACCTGGTGGCGCAGAGCCTCTTCGGTACCTGGAGCACCGGCGGGTGGCGG 249
Db 187 AGCGCACACCTAGTGGCACAGAGCCTCTTCGCTTACCTGGAGCACCGAAGGCTGGCAGC 246
QY 250 GCGGCGCG-----GGGGCGGCTGGATGACGCCACCGCGCGCATGTGGGG 294
Db 247 GCTGCGCGCGCTCCTTTGGCGAAGGGGCGCCCTGGATGCGGCACCTGCACGACGCTGGCA 306
QY 295 CTGACCATCTCGCCTTACAGAGGACCCCGCTACCTGCGGCTGCTGCTGCGCTCGGCG 354
Db 307 CTCACCATCTCAGCCTTACAGAGGATCCCGTTACCTGCGGCTGCTTGAACCTCGCGG 366
QY 355 CGCGCCCTGCTGTATCCGCGCGCGGCTGCGGCTCTCATGCTGGTGGATGGCAACCGC 414
Db 367 CGCGCTTCTGTATCCCGCACACGAGGTTACGCGTGTCTCATGCTGGTGGACGCAACCGC 426
QY 415 GCCGAGGACCTTACATGCTGCAATGTTTCGCGAGGCTTTCGCTGACGAGAGACCCCGCC 474
Db 427 GCTGAGGATCTGTACATGCTGACATGTTCCGAGAAGTCTTCGCGCATGAGGACCCCGC 486
QY 475 ACGTACGTGTGGGCGGCAACTACACAGCCCTCGGAAACCGCGCGCG--CGGCGCG 531
Db 487 ACTTATGTGGGATGGCAACTACCATCAGCCCTGGGAAACAGCGAGGCTACAGGGCGCT 546
QY 532 GTGGCGCGCGGACCTATCGGAGGTGGAGCGGAGGATCTCTGGCGGCTGGCAGTGGAG 591
Db 547 GTCGCTGAAGTGCCTACCGGAGGTGGAGCGGAGGACCCCGGCGGTTGGCGGTGGAG 606
QY 592 GCGCTGGTGAAGACTTCGAGGTGCTGCTGGCGCAGCGCTGGGGCGGCAAGCGCGAG 651
Db 607 GCGCTGGTGAAGACACGCGAGGTGCTGCTGGCTGCTCAGCGCTGGGGCGGCAAGCGTGA 666
QY 652 GTCATGTACAGCCTTCAAGCGGCTCGAGATTGCGGTGGACTACGTGCAGGTCTGTGAC 711
Db 667 GTTATGTACAGCCTTCAAGCGACTGGGCGACTCCGTTGGACTACGTGAGGTCTGTGAC 726
QY 712 TCGGACACAAGGTTGGACCCCATGGCACTGTGGAGCTCGTGGGCTACTGGACGAGGAC 771
Db 727 TCAGACACAAGACTAGACCCCATGGCACTGTGGAGCTTGTGGAGTGTGGATGAAGAC 786
QY 772 CCGCGGTAGGGGCTGTGGTGGGACCTGGGATCTTAAACCTCTGAGACTCTGGGTC 831
Db 787 CCGCGGTAGGGGCTGTGGAGGGGATGTGAGGATCTTAAACCTCTGAGACTCTGGGTC 846
QY 832 AGCTTCTTAAGCAGCCTCGGATCTGCGTACGCTTCAATGTGAGCGGGCTTGTGAGAGC 891
Db 847 AGCTTCTTAAGCAGCTTTCGATACGCTTCAATGTGAGAGAGCTTGTGAGAGC 906
QY 892 TACTTCCACTGTGTATCCTGTCATCAGCGGTCCTCTAG 928
Db 907 TACTTCCACTGTGTCTCCTGTCATCAGTGTCTCTGG 943

AAT91655
ID AAT91655 standard; DNA; 2102 BP.
XX
AC AAT91655;
XX
DT 17-OCT-2003 (revised)
DT 19-DEC-1997 (first entry)
XX
DE Mouse hyaluronate synthase genomic DNA.
XX
KW Hyaluronate synthase; hyaluronic acid; polysaccharide; ds.
XX
OS Mus sp; (strain FM3A).
XX
FH Key Location/Qualifiers
FT CDS 49..1800
FT /tag= a
FT /product= "Hyaluronate_synthase"
XX
PN JP09224674-A.
XX
PD 02-SEP-1997.
XX
PF 26-FEB-1996; 96JP-00038336.
XX
PR 26-FEB-1996; 96JP-00038336.
XX
PA (KAGG) KAGAKU GIJUTSUUCHO CHOKAN KANBO.
XX
DR WPI; 1997-484102/45.
DR P-PSDB; AAW30704.
XX
XX Hyaluronate synthase isolated from mouse cells - useful for large-scale
PT production of hyaluronic acid.
XX
PS Claim 2; Page 10-13; 15pp; Japanese.
XX
CC A mouse-derived cell strain, FM3A P15-A, was cultured in Eagle minimal
CC medium containing 10 % heat-inactivated bovine serum, twice concentration
CC of amino acids and vitamins and penicillin and streptomycin at 37 degrees
CC Celsius. The culture was subjected to immobilised erythrocyte exclusion
CC to examine the extent of extracellular formation of hyaluronic acid
CC matrix. Cells which showed high formation were recovered and named FM3A
CC HAI. Cells synthesising no hyaluronic acid were recovered and named HAS-
CC The HAS- cell in which polyoma large T antigen was expressed was prepared
CC (HAS- P cell). Poly(A) + RNA was isolated from FM3A HAI and cDNA was
CC prepared and was used for constructing a library in HAS- P cells. Cells
CC having hyaluronic acid synthetic activity were selected from the
CC transformants and plasmid DNA was recovered and amplified in E.coli. The
CC resulting genomic DNA sequence codes for hyaluronate synthase having a
CC sequence of 583 amino acids which is used for large-scale production of
CC hyaluronic acid, a component of drugs and cosmetics. (Updated on 17-OCT-
CC 2003 to standardise OS field)
XX
SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;
XX
Query Match 62.6%; Score 677.8; DB 2; Length 2102;
Best Local Similarity 84.5%; Pred. No. 1.4e-107;
Matches 792; Conservative 0; Mismatches 127; Indels 18; Gaps 2;
OY 10 CAGGACGGCGCCCAAGCCACTCTCTGACGGCGCGCTCTGGGCTGGCGGAGGTG 69
DB 55 CAGGACATGCCAAAGCCCTCAGAGCAGCGCGTCTGCTCTGGCTGGCGGCGAGCA 114
OY 70 CTGACCATCGCTTCGCCCTGCTCATCTTGGGCTCATGACCTGGGCTACGCCCGCGGG 129
DB 115 CTCAGATCATCTTGGCTGCTCATCTTGGGCTCATGACCTGGGCTACGCCCGAGGC 174
OY 130 GTGGCGCTGGCTCGCATCGCTACGGCTCTCTGGGCTTGGGCTTACGGGGCTTCCTT 189
DB 175 GTTCCTCTGGCTTCAGATCGCTATGGAATCTCGGCTTTGGCTCTATGGGCGATTCCTC 234
OY 190 TCAGGGCACCTGGTGGCGCAGAGCCTCTTCGGTACCTGGAGCACCGCGGGTGGCGGC 249

235 AGCGCACACCTAGTGGGCACAGAGCCTCTTCCTTACCTGGAGCACCGAAGGTTGGCAGCG 294
250 GCGGCGCG-----GCGGCGCGTGGATGCAGCCACGCGCGCAGTGTGGCG 294
295 GCTGCGCGGCTCTCTTGGCGAAGGGGCGCTGGATGGCGCACTGCACGACGCTGGCA 354
295 CTGACCATCTCGCTTACCGAGGAGGACCGCGGCTACCTGGCCAGTGCCTGGCTGCCGC 354
355 CTCACCATCTCAGCCTACCAAGAGGATCCCGCTTACCTGGCGCCAGTGTTCACCTCCGCG 414
355 CGCGCTCTGTATCCCGCGCGCTGGCGCTGCTTCCTCATGTGGTGGATGGCAACCGC 414
415 CGCGCTTGTGTATCCCGCACACAGAGTTACGCTGCTCATGTGGTGGAGCGGCAACCGC 474
415 CGCGAGGACCTCTATCATGTGTGACATGTTCCGCGAGGTCTTCGTGACGAGGACCCCGCC 474
475 GCTGAGGATCTGTATGTTGGACATGTTCCGAGAGGTCTTCGCGGATGAGGACCCCGCC 534
475 ACCTAGTGTGGAGGCGCACTACACAGCCCTGGGNAACCGCGCGCG-----CGGGCGCG 531
535 ACTTATGTGTGGATGGCAACTACCATCAGCCCTGGGAACCCAGCGGAGGCTACCGGGCGCT 594
532 GTGGCGCGCGAGCCTATCGGAGGTGGAGCGGAGGATCTTGGCGCGCTGGCAGTGGAG 591
595 GTCGGTGAAGTGCTTACCGGAGGTGGAGCGGAGGACCCCGCGCGGTTGGCGTGGAG 654
592 GCGTGTGTGAGGACTCGCAGGTGCTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGCGAG 651
655 GCGTGTGTGAGAACACGCGAGGTGCTGTGCTGCTCAGCGTTGGGGCGGCAACGCTGAG 714
652 GTCATGTACACAGCCTTCAGGCGCTCGGAGTTCCGGTGGACTACGTGCGAGTCTGTGAC 711
715 GTCATGTACACAGCTTTCAAGGGCACTGGGCGACTCCGTGGGACTACGTGCGAGTCTGTGAC 774
712 TCGCACACAAGGTTGGACCCCATCGCACTGCTGGAGCTCGTGGCGGTACTGCGACGAGGAC 771
775 TCAGACACAAGACTAGACCCCATGCGCACTGCTGGAGCTTGGCGAGTGTGGATGAAGAC 834
772 CCGCGGTGAGGGGCTGTTGGTGGGAGCTGCGGATCTTAAACCTCTGGACTCTCTGGGTC 831
835 CCGCGGTGAGGGGCTGTTGGAGGGGATGTGAGGATCTTAAACCTCTGGACTCTCTGGGTC 894
832 AGCTTCTTAGACCGCTTCGATATCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 891
895 AGCTTCTTAGAGCAGTCTTCGATATCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 891
892 TACTTCCACTGTGTATCTTCATCAGCGGCTCCTCTAG 928
955 TACTTCCACTGTGTCTCTGATCAGTGGTCTCTGG 991

RESULT 11
AAZ10862
ID AAZ10862 standard; DNA; 2102 BP.
XX
AC AAZ10862;
XX
DT 19-OCT-1999 (first entry)
XX
DE Hyaluronate synthase coding sequence.
XX
KW Hyaluronate synthase; promoter; mouse; cell growth inhibitor; ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 49..1800
FT /tag= a
XX
PN JP11196875-A.
XX
PD 27-JUL-1999.


```
QY 10 CAGGACGGCCCAAGCCCACTCTCTGACGCCCGCGCTCTCGGCTTGCGCCGCGAGGGTG 69
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 55 CAGGACATGCGCAAGCCCTCAGAGGAGCGCGTCTGCTCTGGCTTGCCAGGCGAGCA 114
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 70 CTGACCATGCGCTTCTGCGCTGCTCATCTCTGGGCTCATGACCTTGCGGCTTACGCCCGCGG 129
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 115 CTGACGATCATCTTCTGCGCTGCTCATCTCTGGGCTCATGACCTTGCGGCTTACGCCCGAGGC 174
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 130 GTGCGGCTGCGCTCGATGCTACGGCTCTCTGCGCTTCTGCGCTTACGGGCGCTTCTT 189
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 175 GTTCTCTGCGCTTCAGATGCTGATGCTCTGGGCTTCTGCGCTTCTGCGGCTTCTCTC 234
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 190 TCAGGCGCACTGCTGCGCAGAGCTCTCTGCGCTTCTGCGCTTCTGCGGCTTCTCTC 249
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 235 AGCGCACACCTTAGTGGCACAGAGCTCTCTGCGCTTCTGCGCTTCTGCGGCTTCTCTC 294
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 250 GCGGCGCG-----GCGGCGCGTGTGATGACAGCCACCGCGCGCAGTGTGGCG 294
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 295 CTGACCATCTCGCTTACGAGGAGCAGCGCGCTTACCTGCGGCTTGCGCTTGCGGCTTGCGG 354
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 355 CTCACCATCTCAGCTTACGAGGATCCCGCTTACCTGCGGCTTGCGCTTGCGGCTTGCGG 414
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 355 CGCGGCTGCTGTACCGCGCGCGGCTGCGGCTCTCATGCTGTGTGATGGCAACCGC 414
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 415 CGCGCTTGTGTACCGCGCACAGAGTTACGCTGCTCATGCTGTGTGATGGCAACCGC 474
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 415 GCGGAGGACCTTACATGCTGCATGTTTCCGAGGTTCTCGCTGACGAGGACCCCGC 474
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 475 GCTGAGGATCTGTACATGCTGACATGTTCCGAGAAGTCTTCCGCGATGAGGACCCCGC 534
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 475 AGGTACGTGTGGAGCGCACTACACAGCCCTGGGAACCGCGCGCG-----GCGGCGG 531
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 535 ACTTATGTGTGGATGGAACCTACCATCAGCCCTGGGAACCAAGCGGAGGCTACGCGCGCT 594
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 532 GTGGCGCGGAGCTTATCGGAGGTGAGGCGGAGGATCTTGGGCGCTGGCAGTGGAG 591
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 595 GTCGTTGAAGTGCCTTACCGGAGGTGAGGCGGAGGAGCCCGCGGCTTGGCGTGGAG 654
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 592 GCGCTGTGTGAGAACCTCGCAGGTGCTGTGCGGCGCAGCGCTGGGCGGCAAGCGCAG 651
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 655 GCGCTGTGTGAGAACCGCAGGTGCTGTGCGTGGCTCAGCGTTGGGCGGCAAGCGTGA 714
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 652 GTCATGTATACAGCTTCAAGCGCTCGGAGATTGCTGGAATTCGTTGAGTCTGAGTCTGTGAC 711
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 715 GTCATGTATACAGCTTCAAGGCACTGGGCGGACTCGGTGGAATTCGTTGAGTCTGTGAC 774
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 712 TCGGACACAGGTTGGACCCCATGGCACTGTGGAGCTCGTGGGCTTACTGGACGAGGAC 771
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 775 TCAGACACAGATAGACCCCATGGCACTGTGGAGCTTGTGGAGTGTGGATGAAGAC 834
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 772 CCGCGGTAGGGCTGTGTGGGAGCGTGGGATCTTAAACCTCTTGAATCTCTGGGTC 831
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 835 CCGCGGTAGGGCTGTGTGGGAGGTGTGAGGATCTTAAACCTCTTGAATCTCTGGGTC 894
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 832 AGCTTCTTACGAGCTGCGTACTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 891
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 895 AGCTTCTTGAAGCAGTCTTCTGATCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 954
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 892 TACTTCCACTGTGATCTCTGATCAGCGTCTCTAG 928
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 955 TACTTCCACTGTGTCTCTGATCAGTGTCTCTG 991
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 13

AAA39987

ID AAA39987 standard; cDNA; 2102 BP.

XX

AC AAA39987;

XX

DT 19-OCT-2000 (first entry)

XX

```
DE Murine HAS1 cDNA.
XX
KW Murine; HAS1; hyaluronate synthase; marker gene; intracellular loop;
KW gene targeting; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 49..1800
FT /*tag= a
FT /product= "HAS1"
XX
PN JP2000116382-A.
XX
PD 25-APR-2000.
XX
PF 13-OCT-1998; 98JP-00291201.
XX
PR 13-OCT-1998; 98JP-00291201.
XX
PA (SEBK ) SEIKAGAKU KOGYO CO LTD.
XX
XX WPI; 2000-369404/32.
XX P-PSDB; AAB09948.
XX
XX DNA used for gene targeting of the hyaluronate synthase gene (HAS1).
XX
XX Example; Page 7-9; 14pp; Japanese.
XX
XX This invention describes a novel DNA which contains an heterologous DNA
XX (I) to be introduced to chromosomal DNA (II) connected respectively to the 5' and 3'
XX second homologous region DNA (II) connected respectively to the 5' and 3'
XX sides of the introduced DNA, and a negative marker gene (III) expressable
XX in the host cell. (I) contains a positive marker gene expressable in the
XX host cell. (I), (II) and the region encoding the intracellular loop of
XX the hyaluronate synthase (HAS) 1 protein from chromosomal DNA is
XX constituted so as to cause a homologous recombination. The DNA can be
XX used for gene targeting of the HAS1 gene. This sequence encodes the
XX murine HAS1 protein described in the method of the invention
XX
SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;
Query Match 62.6%; Score 677.8; DB 3; Length 2102;
Best Local Similarity 84.5%; Pred. No. 1.4e-107;
Matches 792; Conservative 0; Mismatches 127; Indels 18; Gaps 2;
QY 10 CAGGACGGCCCAAGCCCACTCTCTGACGCCCGCGCTCTCGGCTTGCGCCGCGAGGGTG 69
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 55 CAGGACATGCGCAAGCCCTCAGAGGAGCGCGTCTGCTCTGGCTTGCCAGGCGAGCA 114
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 70 CTGACCATGCGCTTCTGCGCTGCTCATCTCTGGGCTCATGACCTTGCGGCTTACGCCCGCGG 129
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 115 CTGACGATCATCTTCTGCGCTGCTCATCTCTGGGCTCATGACCTTGCGGCTTACGCCCGAGGC 174
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 130 GTGCGGCTGCGCTCGATGCTACGGCTCTCTGCGCTTCTGCGCTTACGGGCGCTTCTT 189
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 175 GTTCTCTGCGCTTCTGAGTGTGATGCTCTGGGCTTCTGCGCTTCTGCGGCTTCTCTC 234
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 190 TCAGGCGCACTGCTGCGCAGAGCTCTCTGCGCTTCTGCGCTTCTGCGGCTTCTCTC 249
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 235 AGCGCACACCTTAGTGGCACAGAGCTCTCTGCGCTTCTGCGCTTCTGCGGCTTCTCTC 294
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 250 GCGGCGCG-----GCGGCGCGTGTGATGACAGCCACCGCGCGCAGTGTGGCG 294
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 295 CTGACCATCTCGCTTACGAGGAGCAGCGCGCTTACCTGCGGCTTGCGCTTGCGGCTTGCGG 354
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 355 CTCACCATCTCAGCTTACGAGGATCCCGCTTACCTGCGGCTTGCGCTTGCGGCTTGCGG 414
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 355 CGCGGCTGCTGTACCGCGCGCGGCTGCGGCTCTCATGCTGTGTGATGGCAACCGC 414
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 415 CGCGCTTGTGTACCGCGCACAGAGTTACGCTGCTCATGCTGTGTGATGGCAACCGC 474
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 415 GCGGAGGACCTTACATGCTGCATGTTTCCGAGGTTCTCGCTGACGAGGACCCCGC 474
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 475 GCTGAGGATCTGTACATGCTGACATGTTCCGAGAAGTCTTCCGCGATGAGGACCCCGC 534
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 475 AGGTACGTGTGGAGCGCACTACACAGCCCTGGGAACCGCGCGCG-----GCGGCGG 531
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 535 ACTTATGTGTGGATGGAACCTACCATCAGCCCTGGGAACCAAGCGGAGGCTACGCGCGCT 594
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 532 GTGGCGCGGAGCTTATCGGAGGTGAGGCGGAGGATCTTGGGCGCTGGCAGTGGAG 591
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 595 GTCGTTGAAGTGCCTTACCGGAGGTGAGGCGGAGGAGCCCGCGGCTTGGCGTGGAG 654
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 592 GCGCTGTGTGAGAACCTCGCAGGTGCTGTGCGGCGCAGCGCTGGGCGGCAAGCGCAG 651
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 655 GCGCTGTGTGAGAACCGCAGGTGCTGTGCGTGGCTCAGCGTTGGGCGGCAAGCGTGA 714
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 652 GTCATGTATACAGCTTCAAGCGCTCGGAGATTGCTGGAATTCGTTGAGTCTGAGTCTGTGAC 711
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 715 GTCATGTATACAGCTTCAAGGCACTGGGCGGACTCGGTGGAATTCGTTGAGTCTGTGAC 774
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 712 TCGGACACAGGTTGGACCCCATGGCACTGTGGAGCTCGTGGGCTTACTGGACGAGGAC 771
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 775 TCAGACACAGATAGACCCCATGGCACTGTGGAGCTTGTGGAGTGTGGATGAAGAC 834
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 772 CCGCGGTAGGGCTGTGTGGGAGCGTGGGATCTTAAACCTCTTGAATCTCTGGGTC 831
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 835 CCGCGGTAGGGCTGTGTGGAGGAGGTGTGAGGATCTTAAACCTCTTGAATCTCTGGGTC 894
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 832 AGCTTCTTACGAGCTGCGTACTGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 891
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 895 AGCTTCTTGAAGCAGTCTTCTGATCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 954
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 892 TACTTCCACTGTGATCTCTGATCAGCGTCTCTAG 928
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 955 TACTTCCACTGTGTCTCTGATCAGTGTCTCTG 991
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

QY 415 GCCGAGGACCTCTACATGTCGACATGTTCCGCGAGGTCTTCGCTCGAGGAGACCCGCC 474
Db |||||
QY 475 GCTGAGGATCTGTACATGCTGACATGTTCCGAGAAGTCTTCGCGATGAGGACCCGCC 534
Db |||||
QY 475 AGCTACGTGGGAGCGGCACTACACAGCCCTGGGAACCCGCGCGG---CGGCGGG 531
Db |||||
QY 535 ACTTATGTGTGGATGGCACTACCATCAGCCCTGGGAACCCGCGGAGGCTACCGGGCGCT 594
Db |||||
QY 532 GTGGCGCGGAGCCTATCGGAGGTGAGGCGGAGGATCTTCGGCGGCTGGCAGTGGAG 591
Db |||||
QY 595 GTCGCTGAAGTGCCTACCGGAGGTGAGGCGGAGGACCCCGGCGGTGTGGCGTGGAG 654
Db |||||
QY 592 GCGCTGTGTGAGGACTCGCAGGTGCTGTGCGGCGCAGCGCTGGGGCGGCAAGCGCGAG 651
Db |||||
QY 655 GCGCTGTGTGAGAACACGACGAGTGCCTGTGCGTGGCTCAGCGTTCGGGCGGCAAGCGTGA 714
Db |||||
QY 652 GTCATGTACACAGCCTTCAGGCGCTCGGAGATTCGTTGGACTAGTCGAGTCTGTGAC 711
Db |||||
QY 715 GTCATGTACACAGCTTCAAGGCACTGGGCGACTCGGTGGACTACGTGCAGGTCTGTGAC 774
Db |||||
QY 712 TCGGACACAGGTTGACCCCATGCACTGCTGGAGCTGTCGCGGTACTGACGAGGAC 771
Db |||||
QY 775 TCAGACACAGACTAGACCCCATGCACTGCTGGAGCTTGTGCGAGTGTGGATGAAGAC 834
Db |||||
QY 772 CCCGCGGTAGGGGCTGTGTGGTGGGACGCTGCGGATCTTAACCCCTCTGGAACCTCTGGGTC 831
Db |||||
QY 835 CCCGCGGTAGGGGCTGTGTGGAGGGGATGTGAGGATCTTAACCCCTCTGGAACCTCTGGGTC 894
Db |||||
QY 832 AGCTTCCTAAGCAGCTCGGATCTGCGGTAGCTTCAATGTGGAGCGGGCTGTGAGAGC 891
Db |||||
QY 895 AGCTTCCTGAGCAGCTTTCGATCTGCGGTAGCTTCAATGTGGAGCGAGCTTGTGAGAGC 954
Db |||||
QY 892 TACTTCCACTGTATCTGCACTAGCGCTCTCTAG 928
Db |||||
QY 955 TACTTCCACTGTGTCTGCACTAGCGCTCTCTAG 991
Db |||||

RESULT 14
ID ABQ27384
XX ABQ27384 standard; DNA; 662 BP.
AC ABQ27384;
XX
DT 12-JUL-2002 (first entry)
XX
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 13975.
XX
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX
OS Homo sapiens.
XX
PN WO200218632-A2.
XX
PD 07-MAR-2002.
XX
PP 01-SEP-2001; 2001WO-EP010074.
XX
PR 01-SEP-2000; 2000DE-01043826.
XX
PR 05-SEP-2000; 2000DE-01044543.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
XX WPI; 2002-371829/40.
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.
PS This invention describes a novel method for determining the degree of
XX methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 662 BP; 80 A; 80 C; 247 G; 255 T; 0 U; 0 Other;
Query Match 39.0%; Score 422; DB 6; Length 662;
Best Local Similarity 77.3%; Pred. No. 1.3e-63;
Matches 512; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
QY 34 GCAGCCGCCCTGCTCGGCTCGGCGCTGACCATCTCCGCTCCGATCGCTAC 153
Db 1 GTAGTTTGTCTGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTT 60
QY 94 ATCTTGGCCCTCATGACCTGGGCTTACGCGCGGGGTGCGCTCCGATCGCTAC 153
Db 61 ATTTTGGGTTTATGATTGGGTTTACGTCGTCGGGGTGTCTGTTGGTTTGGATCTTAC 120
QY 154 GGCTCTCTGGGCTTTCGGCTCTACGGGCTCTTCCTTTTCAGCGCACCTCGTGGCGAGAGC 213
Db 121 GGTTTTGTGGTTTTCGGTTTTCGGTTTTCGGTTTTCGGTTTTCGGTTTTCGGTTTTCGGTTT 180
QY 214 CTCTTCGCGTACCTGGAGCACCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 273
Db 181 TTTTTCGCGTATTTGGAGTATCGGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY 274 GCCACCGCGCGCAGTGTGGCTGACCATCTCCGCTCCGCTCCGATCGCTACCTG 333
Db 241 GTTATCGCGCGCTAGTGTGGCGTTGATTATTTTCGTTTATTTAGGAGGATTTTCGCGTATTTG 300
QY 334 CGCCAGTGCCTGGCGTCCGCCCGCGCTGCTGTACCGCGCGCGCGCGCGCGCGCGCGCG 393
Db 301 GGTAGTGTGGGCTGCTGCGGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 360
QY 394 ATGTTGTGGATGCAACCGCGCGGAGACCTCTACATGTGTCGATGTTCCGCGAGTTC 453
Db 361 ATGTTGTGGATGCAACCGCGCGGAGACCTCTACATGTGTCGATGTTCCGCGAGTTC 420
QY 454 TTCGCTGACGAGGACCCCGCGCGGAGACCTCTACATGTGTCGATGTTCCGCGAGTTC 513
Db 421 TTCGTTGACGAGGATTTTCGTTTACGTTGTCGTTGTTGTTGTTGTTGTTGTTGTTGTT 480
QY 514 CCGCGCGCGCGCGCGGCTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 573
Db 481 TTCGCGCGCGCGCGCGGCTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
QY 574 GGGCGGCTGGCAGTGGAGGCGCTGGTGAAGGACTCGCAGGTGCGTGTGCGTGGCGAGCGC 633
Db 541 GGGCGGCTGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 634 TGGGCGCGCAAGCGCGAGGTCATGTACACAGCTTCAAGGCGCTCGGAGATTCGGTGGAC 693
Db 601 TGGGCGCGCAAGCGCGAGGTTATGTATATAGTTTTTAAAGGCGTTTCGGAGATTCGGTGGAT 660

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:08:52 ; Search time 4464.41 Seconds
(without alignments)
9233.835 Million cell updates/sec

Title: US-10-672-399-5
Perfect score: 1083
Sequence: 1 atgagacagcagcagcgc.....tccaagtcgtacttcctcgta 1083

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	922.4	85.2	2072	3	CR602106
2	920.8	85.0	1728	9	AY421569
3	822.4	75.9	1058	4	BMS43754
4	795.6	73.5	965	1	AL545447
5	679.4	62.7	1746	9	AY421571
6	496	45.8	896	5	BX371636
7	479.4	44.3	3542	3	AK053726
8	357.4	33.0	650	2	BB65418
9	311.4	28.8	1084	4	BMS44718
10	297.4	27.5	551	4	BG707429
11	297.4	27.5	551	4	BG707608
12	268.2	24.8	508	9	CE377866
13	262.4	24.2	782	4	BI753116
14	259.4	24.0	610	5	BX327795
15	257.8	23.8	894	7	CF239758
16	253.6	23.4	668	8	BH057613
17	239	22.1	774	5	BP677403
18	236.4	21.8	4151	3	AK028582
19	221.8	20.5	895	5	BU911059
20	209	19.3	747	6	CD466132
21	206.8	19.1	504	4	BG016146
22	202.6	18.7	944	6	CA487249
23	193.8	17.9	495	7	CN664556
24	189.6	17.5	739	9	AY421570

25	187.8	17.3	918	7	CF220501
26	184.2	17.0	660	5	BP674137
27	179.2	16.6	739	7	CN304087
28	179.4	16.6	2891	3	AK079729
29	176	16.3	653	7	CV029395
30	174.8	16.1	857	5	BU906272
31	174.8	16.1	866	6	CA988122
32	172.4	15.9	848	6	CA987283
33	171	15.8	809	5	BP693553
34	170.6	15.8	748	5	BP699960
35	165.8	15.3	647	7	CN666988
36	165.4	15.3	781	5	BP693535
37	164.6	15.2	871	5	BU910414
38	163.6	15.1	920	7	CF220396
39	162.8	15.0	746	5	BP705367
40	162.8	15.0	840	5	BU911086
41	162.8	15.0	941	6	CA790329
42	162.6	15.0	853	6	CA988392
43	162.6	15.0	864	5	BU907701
44	162.6	15.0	898	5	BU906701
45	161	14.9	790	5	BP706190

ALIGNMENTS

RESULT 1
LOCUS CR602106 2072 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CSODI027YJ02 of Placenta Cot 25-normalized of Homo sapiens (human).
ACCESSION CR602106
VERSION CR602106.1 GI:50482913
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2072)
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue
2 (bases 1 to 2072)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES
Location/Qualifiers
1..2072
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI027YJ02"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 85.2%; Score 922.4; DB 3; Length 2072;
Best Local Similarity 90.4%; Pred.No. 4.9e-174;
Matches 1038; Conservative 0; Mismatches 36; Indels 74; Gaps 2;
QY 10 CAGACGCGCCCAAGCCCACTCTCGACACCGCGCTCTCGGCTCGCCGCGAGGTG 69
|||||
DB 58 CAGACGCGCCCAAGCCCACTCTCGACACCGCGCTCTCGGCTCGCCGCGAGGTG 117
|||||

595 CGGAGGTGGAGCGGAGGATCTCTGGCGGCTGGCAGTGGAGCGCTGGTGGAGTCTGT 654
610 AGGTGGTGTGGTGGCAGCGCTGGGGCGGCAAGCGGAGGTATGTACACAGCTTTC 669
655 ANGTTGGTGTGGTGGCGCAGCGCTGGGGCGGCAAGCGGAGGTATGTACACAGCTTTC 714
670 AAGCGCTCGGAGATTCGGTGACTACGTGCAGGTCTGTGACTCGGACACAGGTTGGAC 729
715 AAGCGCTCGGAGATTCGGTGACTACGTGCAGGTCTGTGACTCGGACACAGGTTGGAC 774
730 CCCATGGCACTCTCGTGGAGCTCGTGGGGTACTTGGACGAGGCCCGCGGTAGGGCTGTT 789
775 CCCATGGCACTCTCGTGGAGCTCGTGGGGTACTTGGACGAGGCCCGCGGTANGGGCTGTT 834
790 GGTGGGAGCTGGGATCTTAACTCTGGTCTGGTCTCTGGTCTGGTCTTAAAGCAGCTG 849
835 GNGGGGAGCTGGGAACTTAACTCTGGTCTGGTCTGGTCTGGTCTTAAAGCAACCTG 894
850 CGAT-ACTGGGTAGCTTCAAT-GTGGAGCGGGCTTG 884
895 CGAATACTGGGTAGCTTCAATGGTGGAAACGGGCTTG 931

AL545447 965 bp mRNA linear EST 25-MAR-2004
AL545447 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1027YJ02 5-PRIME, mRNA sequence.
AL545447
AL545447.3 GI:45745929
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 965)
Li W.B., Gruber, C., Jessee, J. and Pollayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31267282.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9502.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?&CS0D1027DE01QPI&c=9502.r.
Location/Qualifiers
1. 965
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1027YJ02"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES
source
1. 965
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1027YJ02"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 73.5%; Score 795.6; DB 1; Length 965;
Best Local Similarity 97.4%; Pred. No. fe-148;
Matches 851; Conservative 9; Mismatches 8; Indels 6; Gaps 5;
10 CAGGACGGCGCAAGCCACTCTCGACCGCGCGCTCTCGGCTGGCGGAGGTG 69
|||||

Db 58 CAGGACGGCGCAAGCCACTCTCTGACGCCCGCTGCTCGCGCTTGGCCCGGA-GGTG 116
Qy 70 CTGACCATCGCTTTCGCCCTCTCTATCTGGGCTCATGACTGGGCTATGCGCCCGG 129
Db 117 CTGACCATCGCTTTCGCCCTCTCTATCTGGGCTCATGACTGGGCTATGCGCCCGG 176
Qy 130 GTGCCCTGGGCTCCGATCGCTAGGGCTCTTGGCTTTCGCCCTCTAGCGGCTTCTCTT 189
Db 177 GTGCCCTGGGCTCCGATCGCTAGGGCTCTTGGCTTTCGCCCTCTAGCGGCTTCTCTT 236
Qy 190 TCAGCGCACTTGGTGGCGCAGAGCTCTTTCGCTACTTGGAGCAACCGCGGTGGCGCG 249
Db 237 TCAGCGCACTTGGTGGCGCAGAGCTCTTTCGCTACTTGGAGCAACCGCGGTGGCGCG 296
Qy 250 GCGCGCGGGGCGCTGGATGCAACCGCGCGCAGTGTGGCTGACCATCTCCGCC 309
Db 297 GCGCGCGGGGCGCTGGATGCAACCGCGCGCAGTGTGGCTGACCATCTCCGCC 356
Qy 310 TACCAGGAGGACCCCGGTACTCTCGCCAGTGTGGCTTTCGCCCGCGCTTCTGTGAC 369
Db 357 TACCAGGAGGACCCCGGTACTCTCGCCAGTGTGGCTTTCGCCCGCGCTTCTGTGAC 416
Qy 370 CCGCGCGCGGCTGCGCTCTCATGTGTGGATGCAACCGCGCGCGAGACCTCTTAC 429
Db 417 CCGCGCGCGGCTGCGCTCTCATGTGTGGATGCAACCGCGCGCGAGACCTCTTAC 476
Qy 430 ATGTGACATGTTCCGGGAGGTCTTGTGACAGAGACCCCGGCACTGCTGTGGGAC 489
Db 477 ATGTGACATGTTCCGGGAGGTCTTGTGACAGAGACCCCGGCACTGCTGTGGGAC 536
Qy 490 GGCRAACTACCAAGCCCTGGGAAACCCCGCGCGCGGTGGGCGCGGAGCCCTAT 549
Db 537 GGCRAACTACCAAGCCCTGGGAAACCCCGCGCGCGGTGGGCGCGGAGCCCTAT 596
Qy 550 CCGGAGGTGGAGGCGGAGATCTTGGCGGCTGGCAGTGGAGGCGCTGGTGAAGACTCGC 609
Db 597 CCGGAGGTGGAGGCGGAGATCTTGGCGGCTGGCAGTGGAGGCGCTGGTGAAGACTCGC 656
Qy 610 AGTGTGCTGTGGTGGCGCAGCGCTGGGCGGCAAGCGCGGAGTCTGTATACACAGCTTC 669
Db 657 AGTGTGCTGTGGTGGCGCAGCGCTGGGCGGCAAGCGCGGAGTCTGTATACACAGCTTC 716
Qy 670 AAGCGCTCGGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACACAAAGTGGAC 729
Db 717 AAGCGCTCGGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACACAAAGTGGAC 776
Qy 730 CCCATGGCACTCTGGAGCTGTGGGCTACTGGACGAGGACCCCGGGTAGGGCTGTT 789
Db 777 CCCATGGCACTCTGGAGCTGTGGGCTACTGGACGAGGACCCCGGGTAGGGCTGTT 836
Qy 790 GGTGGGACGTGGGATCTTAACTCTGGATCTTGGCTTCTGGCTCAGCTTCTTAAAGCAGCTG 849
Db 837 GGTGGGACGTGGGATCTTAACTCTGGATCTTGGCTTCTGGCTCAGCTTCTTAAAGCAGCTG 892
Qy 850 CGATACTGGGTAGCTTCAATGTGGAGCGGCTT 883
Db 893 CGATACTGGGTAGCTTCAATGTGGAGCGGCTT 925

RESULT 5
LOCUS AY421571 1746 bp DNA linear GSS 17-DEC-2003
DEFINITION Mus musculus Hs1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY421571
VERSION AY421571.1 GI:39748430
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1746)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,

Db 836 GCCAGTGGCTGGGTCGCGCCCGCCCTTGCTGTATACCGGGGGCGGTGGGGCGTCTCTCA 837

Qy 335 TGGTGGTGGATGGCAACCGCGCCGAGGACCTCTTAATGGTGCACATGTTCCGCGAGGTCT 454

Db 836 TGGT-GTGGAGGCAACCGCGCCGAGGACCTTTACAAATGTTGCACA-GTTTCCCGAGGTCT 779

Qy 455 TCGCTGACGAGACCCCGCCAGTACGTGTGGAGCGCAACTTACCACACAGCCCTGGG-AA 513

Db 778 TCGCTGACGAGACCCCGCCAGTACGTGT-GGACGGCAACTTACCACACAGCCCTGGGAAA 720

Qy 514 CCGCGCGCGCGCGCGCGTGGCGCCCGAGGCTTATCGGGAGGTGGAGCGCGAGGATCCT 573

Db 719 CCNCGNCGNCGGCGCGGTGGCNCGGAGCTTATCGGAGGTGAGNCGNAGGATCCT 660

Qy 574 GGGCGGCTGGCAGTGGAGCGCTGTGTAGGACTTCCAGAGTGGTGGTGGCGGAGCGC 633

Db 659 GGNCGGCTGGCAGTGGAGCGCTGTGTAGGACTTCCAGAGTGGTGGTGGCGGAGCGC 600

Qy 634 TGGGGCGGACGCGAGGTCATGTACACAGCTTCAAGGGCGCTCGGAGATTCGGTGGAC 693

Db 599 TGGGGCGGACGCGAGGTCATGTACACAGCTTCAAGGGCGCTCGGAGATTCGGTGGAC 540

Qy 694 TACGTGAGGTCTGTGACTCGGACACAAAGGTTGGACCCCATGGCACTGCTGGAGTCTCGT 753

Db 539 TACGTGAGGTCTGTGACTCGGACACAAAGTGGACCCCATGGCACTGCTGGAGTCTCGT 480

Qy 754 CGGTGACTGGACGAGGACCCCGGTAGGGCTGTGTGGGAGCGTGGGATCTTAAAC 813

Db 479 CGGTGACTGGACGAGGACCCCGGTAGGGCTGTGTGGGAGCGTGGGATCTTAAAC 420

Qy 814 CCTCTGAGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 873

Db 419 CCTCTGAGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360

Qy 874 GAGCGGGCTGTGAGAGTACTTCCACGTGTGTATCTCGCATCAGCGGTCCTCTAG----- 928

Db 359 GAGCGGGCTGTGAGAGTACTTCCACGTGTGTATCTCGCATCAGCGGTCCTCTAGGGCTA 300

Qy 929 -----AATCTCGGCC 938

Db 299 TATAGGAATAACCTCTTGAGCAGTTCCTTGAGGCTGGTACACAGAGTTCCTGGGT 240

Qy 939 AGGCCCCAGGAGCAGCGGATGATGCTCTATCTCTCGCCCC----- 980

Db 239 ACCCACTGTATTTGGGGATGACCGGCACCTTCAACACCGATGCTCAGCATGGGTAT 180

Qy 981 -CGTGACGTACACCTCCAGGTCCCGCTGCTACTCAGAGCGCCCTCTCTCTCTCTCTCGG 1039

Db 179 GCTACCAAGTACACCTCCAGGTCCCGCTGCTACTCAGAGCGCCCTCTCTCTCTCTCGG 120

Qy 1040 TGGCTGAGCCAGCAGACACGCTGTGTCGAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1083

Db 119 TGGCTGAGCCAGCAGACACGCTGTGTCGAAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 76

RESULT 7

AK053726

LOCUS

DEFINITION Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:EI30302P16 product:hyaluronan synthase1, full insert sequence.

ACCESSION AK053726

VERSION AK053726.1 GI:26343670

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1

AUTHORS Carninci,P. and Hayashizaki,Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 99279253

10349636

2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

6 (bases 1 to 3542)

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saichou,R., Saichou,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.jp/ Location/Qualifiers 1. 3542 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM_DB:E130302P16" /db_xref="taxon:10090" /clone="E130302P16" /tissue_type="eyeball" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="0 day neonate"

FEATURES

source

CDS	61...456 /note="unnamed protein product; hyaluronan synthase1 (MGDI GI:106590, GB NM_008215, evidence: BLASTN, 99%, match=762) putative" /codon_start=1 /protein_id="BAC35492.1" /db_xref="GI:26343671" /translation="MRQMPKPESEARCCSGLARRALTIFALLIIGLMTWAYAGVP LASDRYLAFGLYGAFLSAHLVAOISLPAYLEHRRVAAAARRSLAKGPLDAAATARSVA LTISAQEDPAYLRQCLTSARALLPHTR"	DEFINITION BB665418 RIKEN full-length enriched, 0 day neonate eyeball Mus musculus cDNA clone E130302P16 5', mRNA sequence. BB665418 VERSION BB665418.1 GI:16499056 EST. Mus musculus (house mouse) Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 650) Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y. RIKEN Mouse ESTs (Arakawa,T., et al. 2001) Unpublished (2001) Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/ Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Kira,A., Matsunura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y. Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y. Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001) Please visit our web site (http://genome.gsc.riken.go.jp) for further details. e mouse tissues Location/Qualifiers 1..650 /organism="Mus musculus" /mol_type="mRNA" /db_xref="taxon:10090" /clone="E130302P16" /tissue_type="eyeball" /dev_stage="0 day neonate" /lab_host="DH10B" /clone_lib="RIKEN full-length enriched, 0 day neonate eyeball" /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGCGCGCCGCACTCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT , cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by
ORIGIN	Query Match 44.3%; Score 479.4; DB 3; Length 3542; Best Local Similarity 81.6%; Pred. No. 1.4e-85; Matches 598; Conservative 0; Mismatches 116; Indels 19; Gaps 3; QY 10 CAGGAGCGCCCAAGCCACTCTCTGAGCGCGCGCTGCTCGGCTGCGCGGAGGTG 69 DB 67 CAGGACATCCCAAGCCCTCAGAGCAGCGCTGTGCTCTGCGCTGCGCGGAGCA 126 QY 70 CTGACCATCGCTTCGCCCTGTCTATCCTGGGCTCATGACCTGGGCTTACGCGCGCGG 129 DB 127 CTCAGATCATTTTGCCCTGTCTATCCTGGGCTCATGCTGGGCTTACGCGCGAGGC 186 QY 130 GTGCGCTGGCTCCGATCGCTACGGCCCTCTGGGCTCTCGGCTCTACGGGCGCTTCCTT 189 DB 187 GTTCTCTGCTTCAGATCGCTATGGACTCTGGCTTTGGCTCTATGGGCGATTCTCTC 246 QY 190 TCAGCGCACTGGTGGCGCAGAGCTCTTCGGGTACTCTGGAGACCGCGGGTGGCGGG 249 DB 247 AGCGCACACATGTGCGCAGAGCTCTTCGCTTACTCTGGAGCACCGAAGGGTGGCAGG 306 QY 250 GCGGCGCG-----GGGCGCGCTGGATGTCAGCGCACCGCGCGCATGTGGCG 294 DB 307 GCTGCGCGCGCTCTTGGCGAAGGGCGCCCTGGATGCGGCGCCTGTCACGCGCGTGGCA 366 QY 295 CTGACCATCTCGGCTTACCGAGGAGACCGCGGCTACTTCGCGCATGCTGCGCTCGCGCC 354 DB 367 CTCACCATCTCAGCTTACCAAGAGATCCCGCTTACTTCGCGCATGCTTGACCTCGCG 426 QY 355 CGCGCGCTCTGTACCGCGCGCGCGCTGCGGCTCTCATGTTGGTGGATGGCAACCGC 414 DB 427 CGCGCTTCTGTGTACCGCGCACAGGTAACCGCTGCTCATGTTGGTGGACGCGCAACCGC 486 QY 415 GCGGAGGACCTCTACATGTTCCGACATGTTCCGCGAGTCTTCGCTGACGAGCACCGCGC 474 DB 487 GCTGAGGATCTGTACATGTTGGACATGTTCCGAGAAGTCTTCGCGATGAGGACCGCGCC 546 QY 475 ACGTAGTGTGGAGCGCAACTACCAACAGCCCTTGGGAACCGCGCGCGG---CGGCGCGG 531 DB 547 ACTTATGTGGATGGCAACTACCATCAGCCCTGGGAACCAAGCGGACTACGGGCGCT 606 QY 532 GTGGGCGCGGAGCTTATCGGAGGTGGAGCGGAGGATCTTGGGCGGTGGCGAGTGGAG 591 DB 607 GTCGGTGAAGGTGCTTACCGGAGGTGGAGGCGGAGACCCCGGCGGTGGCGGTGGAG 666 QY 592 GCGCTGTTGAGGAC-TCCGAGTGGTGTGCGTGGCGGAGCGCTGGGCGGCAACGCGCA 650 DB 667 GCGCTGTTGAGAACAGCGAGGTGCGTGTGCGTGTGCGTGTGCGGCGGCAACGCTGA 726 QY 651 GGTCTATGTACACAGCTTCAAGGCGCTCGGAGATTCGGTGGGACTAGTGCAGGTCTGTGA 710 DB 727 GGTCTATGTACAGCTTTCAGGCACTGGGCGCACTCCGTGGAGTACGTGCAGGTGAGTGA 786 QY 711 CTCGACACAAAGG 723 DB 787 TGGGTTTCCAGG 799	FEATURES source
LOCUS	BB665418 650 bp mRNA linear EST 26-OCT-2001	


```

VERSION      BG707429.1  GI:13983769
KEYWORDS
SOURCE       Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE        National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
COMMENT      Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
              cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
              Toshiyuki and Piero Carninci (RIKEN)
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM10678 row: p column: 12
              High quality sequence stop: 551.
              Location/Qualifiers
                1..551
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:4795643"
                  /tissue_type="hypothalamus"
                  /lab_host="DH10B"
                  /clone_lib="NIH_MGC_96"
                  /note="Organ: brain; Vector: pBluescriptR (modified
                  pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
                  (gtcgag); Oligo-dt primed using primer
                  5'-TTTTTTTTTTTTTTVN-3', size-selected for average
                  insert size 2.3 kb and normalized to ROT 5. This is a
                  primary library enriched for full-length clones and
                  constructed using the Cap-trapper method (Carninci, in
                  preparation). Library constructed by M. Brownstein
                  (NIH/NHGRI, National Institutes of Health). Note: this is
                  a NIH_MGC Library."

FEATURES
    source
        1..551
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4793704"
            /tissue_type="hypothalamus"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_96"
            /note="Organ: brain; Vector: pBluescriptR (modified
            pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
            (gtcgag); Oligo-dt primed using primer
            5'-TTTTTTTTTTTTTTVN-3', size-selected for average
            insert size 2.3 kb and normalized to ROT 5. This is a
            primary library enriched for full-length clones and
            constructed using the Cap-trapper method (Carninci, in
            preparation). Library constructed by M. Brownstein
            (NIH/NHGRI, National Institutes of Health). Note: this is
            a NIH_MGC Library."

ORIGIN
Query Match      27.5%; Score 297.4; DB 4; Length 551;
Best Local Similarity 79.0%; Pred. No. 3.3e-49;
Matches 413; Conservative 0; Mismatches 36; Indels 74; Gaps 2;

QY      635  GGGGGGCGAAGCGGAGGTCATGTATACACAGCCCTTCAAGCGCTCGGAGATTCGGTGGACT 694
        |||||
        6    GGGGGGCGAAGCGGAGGTCATGTATACACAGCCCTTCAAGCGCTCGGAGATTCGGTGGACT 65
QY      695  ACGTCAGTCTGTGACTCGGACACAAAGTTGGACCCATGACCTCGTGGAGCTCGTGC 754
        |||||
        66  ACGTCAGTCTGTGACTCGGACACAAAGTTGGACCCATGACCTCGTGGAGCTCGTGC 125
QY      755  GGGTACTGGACAGGACCCCGGGTAGGGGCTGTTGGTGGGACCTGGCGATCCTTAACC 814
        |||||
        126  GGGTACTGGACAGGACCCCGGGTAGGGGCTGTTGGTGGGACCTGGCGATCCTTAACC 185
QY      815  CTCTGGACTCTCGGTGAGCTTCCTTAAGCAGCGCTCGGATACATGGGTAGCTTCAATGCG 874
        |||||
        186  CTCTGGACTCTCGGTGAGCTTCCTTAAGCAGCGCTCGGATACATGGGTAGCTTCAATGCG 245
QY      875  AGCGGGCTTGTGACAGACTTCCCACTGTATCTCGATCAGCGGTCTCTAG----- 928
        |||||
        246  AGCGGGCTTGTGACAGACTTCCCACTGTATCTCGATCAGCGGTCTCTAGGGCCCTAT 305
QY      929  -----AATCCTGCCCA 939
        |||||
        306  ATAGGAATAACCTCTTGCAGCAGTTCTTGTAGGCCCTGGTACACAGAGTTCTCTGGGTA 365
        |||||
        940  GGCCCCAGGGAGCA CGGAGATGATGCCCTCATTTCTTCGCCCC----- 980
        |||||

```

```

Db      366  CCCACTGTACTTTTGGGGATGACCGGCACCTCCACAAACGGCATGCTAGCATGGGTATG 425
QY      981  CGTCAGGTACACCTCCAGGTCCTGCTACTCTAGAGACGCCCTCGTCTCTTCTCGCGGT 1040
        |||||
        426  CTACCAAGTACACCTCCAGGTCCTGCTACTCTAGAGACGCCCTCGTCTCTTCTCGCGGT 485
        |||||
        1041  GGCTAGGCGCAGCAGACACGCTGGTCCAAAGTCGTACTTCCGTGA 1083
        |||||
        486  GGCTAGGCGCAGCAGACACGCTGGTCCAAAGTCGTACTTCCGTGA 528
        |||||

RESULT 11
BG707608
LOCUS
DEFINITION  BG707608 551 bp mRNA linear EST 07-MAY-2001
            mRNA sequence.
ACCESSION  BG707608
VERSION    BG707608.1 GI:13984125
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 551)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
            Toshiyuki and Piero Carninci (RIKEN)
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10673 row: o column: 17
            High quality sequence stop: 551.
            Location/Qualifiers
              1..551
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4793704"
                /tissue_type="hypothalamus"
                /lab_host="DH10B"
                /clone_lib="NIH_MGC_96"
                /note="Organ: brain; Vector: pBluescriptR (modified
                pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
                (gtcgag); Oligo-dt primed using primer
                5'-TTTTTTTTTTTTTTVN-3', size-selected for average
                insert size 2.3 kb and normalized to ROT 5. This is a
                primary library enriched for full-length clones and
                constructed using the Cap-trapper method (Carninci, in
                preparation). Library constructed by M. Brownstein
                (NIH/NHGRI, National Institutes of Health). Note: this is
                a NIH_MGC Library."

ORIGIN
Query Match      27.5%; Score 297.4; DB 4; Length 551;
Best Local Similarity 79.0%; Pred. No. 3.3e-49;
Matches 413; Conservative 0; Mismatches 36; Indels 74; Gaps 2;

QY      635  GGGGGGCGAAGCGGAGGTCATGTATACACAGCCCTTCAAGCGCTCGGAGATTCGGTGGACT 694
        |||||
        6    GGGGGGCGAAGCGGAGGTCATGTATACACAGCCCTTCAAGCGCTCGGAGATTCGGTGGACT 65
QY      695  ACGTCAGTCTGTGACTCGGACACAAAGTTGGACCCATGACCTCGTGGAGCTCGTGC 754
        |||||
        66  ACGTCAGTCTGTGACTCGGACACAAAGTTGGACCCATGACCTCGTGGAGCTCGTGC 125
QY      755  GGGTACTGGACAGGACCCCGGGTAGGGGCTGTTGGTGGGACCTGGCGATCCTTAACC 814
        |||||
        126  GGGTACTGGACAGGACCCCGGGTAGGGGCTGTTGGTGGGACCTGGCGATCCTTAACC 185
QY      815  CTCTGGACTCTCGGTGAGCTTCCTTAAGCAGCGCTCGGATACATGGGTAGCTTCAATGCG 874
        |||||
        186  CTCTGGACTCTCGGTGAGCTTCCTTAAGCAGCGCTCGGATACATGGGTAGCTTCAATGCG 245
QY      875  AGCGGGCTTGTGACAGACTTCCCACTGTATCTCGATCAGCGGTCTCTAG----- 928
        |||||
        246  AGCGGGCTTGTGACAGACTTCCCACTGTATCTCGATCAGCGGTCTCTAGGGCCCTAT 305
QY      929  -----AATCCTGCCCA 939
        |||||
        306  ATAGGAATAACCTCTTGCAGCAGTTCTTGTAGGCCCTGGTACACAGAGTTCTCTGGGTA 365
        |||||
        940  GGCCCCAGGGAGCA CGGAGATGATGCCCTCATTTCTTCGCCCC----- 980
        |||||

```

```
Db 126 GGGTACTGGACGAGACCCCGGGTAGGGCTGTTGGTGGGACGTGCGGATCCTTAACC 185
QY 815 CTCTGGACTCTCGGTGACGTCTTCAAGCAGCCTCGGATACGTGGTAGCCCTTCAATGCG 874
Db 186 CTCTGGACTCTCGGTGACGTCTTCAAGCAGCCTCGGATACGTGGTAGCCCTTCAATGCG 245
QY 875 AGCGGGCTTGTGACAGACTTCTCACTGTGTATCTCGCATCAGCGGTCTCTAG----- 928
Db 246 AGCGGGCTTGTGACAGACTTCTCACTGTGTATCTCGCATCAGCGGTCTCTAGCGCCTAT 305
QY 929 -----AATCTGTGCCCA 939
Db 306 ATAGGAATACTCTTGCAGCAGTTCTTGGAGCCTGGTACAAACAGAACTTCTGGGTA 365
QY 940 GGCCCCAGGAGCAGCGATGATGCCCTCATCTTCGCGCCC----- 980
Db 366 CCCACTGTACTTTTGGGGATGACCGGCACCTCACCAACCGCATGCTCAGCATGGGTTATG 425
QY 981 CGTCAGGTACACCTCCAGGTCCTCGCTCTACTCAGACGCGCTCGCTTCCTCGGCT 1040
Db 426 CTACCAAGTACACCTCCAGGTCCTCGCTCTACTCAGACGCGCTCGCTTCCTCGGCT 485
QY 1041 GCCTGAGCCAGCAGACGCTGGTCCAAAGTCGTACTTCCGTGA 1083
Db 486 GCCTGAGCCAGCAGACGCTGGTCCAAAGTCGTACTTCCGTGA 528

RESULT 12
CE377866
LOCUS tigr-gas-dog-17000362008906 Dog Library Canis familiaris genomic,
DEFINITION genomic survey sequence.
ACCESSION CE377866
VERSION CE377866.1 GI:36605926
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 508)
AUTHORS Kirkness,B.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
MEDLINE 22875432
PUBMED 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
FEATURES             Location/Qualifiers
     source           1..508
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site 1: BstXI; Libraries were prepared from
                     peripheral blood"

ORIGIN
Query Match          24.8%; Score 268.2; DB 9; Length 508;
Best Local Similarity 91.1%; Pred. No. 2.2e-43;
Matches 285; Conservative 0; Mismatches 28; Indels 0; Gaps 0;
QY 397 GTGTGGATGCCAACCGCGCAGGACCTCTACATGGTCGACATGTTCCGCGAGGTCTTC 456
|||||
```

```
Db 1 GTGTGGACGCGCAACCGCCCGAGCACTCTACATGTTGGACATGTTCCGCGAGGTCTTC 60
QY 457 GCTGACGAGACCCCGCCACGTACTGTGGGACGGCACTACCAACCGCCCTGGGAACCC 516
Db 61 GCCACGAGGACCCCGGGCACTACGTGTGGGACGGCAACTACCAACCGCCCTGGGAGCCG 120
QY 517 GCGCGCGCGCGCGGTGGCGCGCGAGCCTATCCGAGAGTGGAGGCGGAGGATCCTGGG 576
Db 121 CGCGCGCGCGCGCGGAGGCGCGGCGCTTACCGGAGGTGGAGGCGCGAAGACCCCGGG 180
QY 577 CGCTGCGAGTGGAGGCGCTGGTGAGGACTCGCAGGTCGCTGTCGTCGTCGTCGTCGTCG 636
Db 181 CGCTGCGGTGGAGGCGCTGGTGAGGACGACAGGTGCGTGTGCGTGGCGCAGCGCTGG 240
QY 637 GGCGGCAAGCGCGAGGTATGTACACAGCCTTCAAGCGGCTCGAGATTCGGTGGACTAC 696
Db 241 GGCGGCAAGCGCGAGGTATGTACACAGCCTTCAAGCGGCTCGATGATCTCTGTGGACTAC 300
QY 697 GTGCAGGTCTGTG 709
Db 301 GTGCAGGTGAGTG 313

RESULT 13
BI753116
LOCUS BI753116
DEFINITION 603025961F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196399 5',
mRNA sequence.
ACCESSION BI753116
VERSION BI753116
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 782)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-x@mail.nih.gov
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11491 row: j column: 16
High quality sequence stop: 780.
FEATURES             Location/Qualifiers
     source           1..782
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5196399"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_114"
                     /note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI;
                     Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is destroyed
                     upon cloning). Average insert size 1.5 kb, insert size
                     range 1-3 kb. Library is normalized and enriched for
                     full-length clones and was constructed by C. Gruber
                     (Invitrogen). Research Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."

ORIGIN
Query Match          24.2%; Score 262.4; DB 4; Length 782;
Best Local Similarity 77.6%; Pred. No. 3.2e-42;
Matches 377; Conservative 0; Mismatches 36; Indels 73; Gaps 2;
```

```
QY 671 AGGCGCTCGGAGATTGCGTGGAGTCTAGCTGCGAGGTCTGTGACTCGGACACAGGTTGGACC 730
Db 6 AAGCGCTCGGAGATTGCGTGGAGTCTAGCTGCGAGGTCTGTGACTCGGACACAGGTTGGACC 65
QY 731 CCATCGCACTGCTGAGACTCGTGGGTTACTGAGAGGACCCCGGGTAGGGGCTGTG 790
Db 66 CCATCGCACTGCTGAGACTCGTGGGTTACTGAGAGGACCCCGGGTAGGGGCTGTG 125
QY 791 GTGGGAGCTGCGGATCTCTTAACCTCTGGACTCTGAGTCTCTGGGTCAAGTTCCTTAAGCAGCCTGC 850
Db 126 GTGGGAGCTGCGGATCTCTTAACCTCTGGACTCTGAGTCTCTGGGTCAAGTTCCTTAAGCAGCCTGC 185
QY 851 GATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGTCTTCACTGTGTATCT 910
Db 186 GATACCTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGTCTTCACTGTGTATCT 245
QY 911 GCATCAGCGGCTCTCTA----- 927
Db 246 GCATCAGCGGCTCTCTAGGCCCTATATAGGAATAAAGTCTTTGAGGAGCTTCTTGGAGGCT 305
QY 928 GAATCTCTGCCAGGCCCGGAGGACGCGATGATGCCCTCATTTCTCGCCCC----- 980
Db 306 GGTACACACAGAGTTCTTGGGTACCCACTGTACTTTGGGATGACCGGACCTCACCA 365
QY 981 -----CGTGAGGTACACCTCCAGTCCCGCTGCTACTCAGA 1017
Db 366 CGCATGCTCAGCATGGGTTATGTACCAAGTACACCTCCAGTCCCGCTGCTACTCAGA 425
QY 1018 GAGCGCTCTGCTTCTCGGTGCTGAGCCAGCAGACAGCTGCTGTCAGTCTT 1077
Db 426 GAGCGCTCTGCTTCTCGGTGCTGAGCCAGCAGACAGCTGCTGTCAGTCTT 485
QY 1078 CCGTGA 1083
Db 486 CCGTGA 491

RESULT 14
BX327795
LOCUS BX327795 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS01027YJ02 5-PRIME, mRNA sequence.
ACCESSION BX327795
VERSION BX327795.2 GI:46277978
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 610)
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 2, 2003 this sequence version replaced gi:30340686.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9502.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0BAG028ZH06_CS02640_1&c=9502.r

FEATURES
source
1. 610
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01027YJ02"
```

```
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
(note="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized.")

ORIGIN
Query Match 24.0%; Score 259.4; DB 5; Length 610;
Best Local Similarity 97.4%; Pred. No. 1.3e-41;
Matches 263; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 659 ACACAGCCTCAAGCGCTCGAGATTGCGTGGAGTCTAGTGCAGGTCTGTGACTCGGACA 718
Db 1 ACATGCTCTTAAGNCGCTCGGAGATTGCGTGGAGTCTAGTGCAGGTCTGTGACTCGGACA 60
QY 719 CAAGTTGGACCCCATGGCACTGCTGGAGTCTGCGGGTACTGACGAGACCCCGGG 778
Db 61 CAAGTTGGACCCCATGGCACTGCTGGAGTCTGCGGGTACTGACGAGACCCCGGG 120
QY 779 TAGGGGCTGTTGGTGGGACGCTGGGATCTCTTAACCTCTCGAGTCTCTGGGTCAAGTTC 838
Db 121 TAGGGGCTGTTGGTGGGACGCTGGGATCTCTTAACCTCTCGAGTCTCTGGGTCAAGTTC 180
QY 839 TAAGCAGCTCGGATCTGCGGTAGCTTCAATGTGGAGCGGGCTTGTGAGAGTCTTCC 898
Db 181 TAAGCAGCTCGGATCTGCGGTAGCTTCAATGTGGAGCGGGCTTGTGAGAGTCTTCC 240
QY 899 ACTGTGATCTCGATCAGCGGCTCTCTAG 928
Db 241 ACTGTGATCTCGATCAGCGGCTCTCTAG 270

RESULT 15
CF239758
LOCUS CF239758 894 bp mRNA linear EST 05-AUG-2003
DEFINITION AGENCOURT_15100464 NICHD_XGC_Emb6 Xenopus tropicalis cDNA clone
IMAGE:6994934 5', mRNA sequence.
ACCESSION CF239758
VERSION CF239758.1 GI:33442966
KEYWORDS EST.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 894)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgabs-remail.nih.gov
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14677 Row: m Column: 13
High quality sequence stop: 699.
Location/Qualifiers
1. 894
/organism="Xenopus tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="IMAGE:6994934"
/tissue_type="neurula"
/dev_stage="embryo, stages 14-19"
/lab_host="DH10B (phage-resistant)"
```

/clone.lib="NICHX XGC Emb6"
/note="Vector: pCMV-SF0R6.1; Site 1: NotI; Site 2: EcoRV;
Cloned unidirectionally. Primer: Oligo df. Average insert
size 2.1 kb. Constructed by Invitrogen. Note: This is a
Xenopus Gene Collection (XGC) library."

ORIGIN

```
Query Match      23.8%; Score 257.8; DB 7; Length 894;
Best Local Similarity 62.5%; Pred. No. 2.6e-41;
Matches 422; Conservative 0; Mismatches 247; Indels 6; Gaps 1;

Qy 289 GTGGCGCTGACCATCTCCGCTTACAGGAGGACCCCGGTACCTGCGCAGTGCTTGCG 348
Db 54 GTGGCGCTGACCATCTCCGCTTACAGGAGGACCCCGGTACCTGCGCAGTGCTTGCG 113
Qy 349 TCCGCGCCCGCCCTCTCTACCCGCGCGCGGCTGCGCGTCTCTCANGTGGTGGATGCC 408
Db 114 TCCGCGCGCTACGTGAAGTACCCCAAGGACAAATCAAGATCATTTTGGTTCATCGACGG 173
Qy 409 AACCGCGCGGAGGACCTCTACATGGTGCACATGTTCCGCGAGGTCTTCGCTGACGAGGAC 468
Db 174 AACCGGAGGATGACCGCTTACATGATGGAGATGTTCAAAGATGTTCTCCATGGTGAACGAC 233
Qy 469 CCGCGCAGCTACGTGTGGAGCGGAACTACACAGCCCTGCGGAACCCGCGCGCGCGGCG 528
Db 234 GTGGGTACTACTAGCTCTGAAGGGCAACTACACACTGGCGGTAAAGGAGACCCAGATGGC 293
Qy 529 GCGGTGGCGCGCGGAGCCTATCGGAGGTGAGAGCGGAGGATCTCTGGCGCGCTGGCAGTG 588
Db 294 TCCTGTCTCTGAGGTGTCCAAAGCCCTGAATGAAGATGAG-----GGGATACGTATAGTG 347
Qy 589 GAGCGCTGTGTGAGGACTCGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGCGGCGCAAGCGC 648
Db 348 GAGGAACTGGTTCAGAACCAAGAGGTGTGTGTGCATCATGCGCAGTGGGCGGCGCAACGG 407
Qy 649 GAGGTCAATGTACACAGCTTCAAGCGCTTCAGGAGATTCGGTGGACTACGTGCAGGTCTGT 708
Db 408 GAGGTCAATGTACACAGGCAATTCGGGCGCATCGGGACCAACATGGACTACGTACAGGTCTGT 467
Qy 709 GACTCGGACACAAGGTTGGAACCCCATGCGACTGTGGAGCTCGTGGCGGTACTGGACGAG 768
Db 468 GACTCCGACACCAAACTGGACGMACTGGCCACGGTGGAAATGGTGAAGGTCCTAGAGGCC 527
Qy 769 GACCCCGGGTAGGGCTTTGGTGGGACGCTGGGATCCTTAACCCCTCTGGACTCCTGG 828
Db 528 AACGAGCTGTGCGGGCGGTGGGAGGACGCTTGCATCTCTGAACCCCTACGACTCTCTTC 587
Qy 829 GTCAGCTTCTTAAGCAGCTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGCTTGTGAG 888
Db 588 ATCAGTTTCATGAGCAGCCTACGTTATTGGATGGGTTTACGTGGAGAGGCGCTGCCAG 647
Qy 889 AGCTACTTCCACTGTGTATCTTGCATACAGCGGTCTCTAGAAATCTGCCCCAGGCCACAG 948
Db 648 TCCTACTTCGACTGCGTCTCTGTATCAGTGGCCCTCTGGAAATGTACCGGAACGACATT 707
Qy 949 GAGCAGCGATGATG 963
Db 708 CTCAGGTGTTT 722
```

Search completed: March 13, 2005, 11:54:57
Job time : 4466.41 secs

Db 36 ATGAGACGACGAGCGCCCAAGCCCACTCTCTGACGCGCGGCTGCTCGGCGCTGACC 95
Qy 61 CGAGGGTGTGACCATGCGCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
Db 96 CGAGGGTGTGACCATGCGCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 155
Qy 121 GCGCGCGGGTGGCGCTGCGCTCCGATGCTAGCGCTCTGCGCTTGGGCTCTACGGG 180
Db 156 GCGCGCGGGTGGCGCTGCGCTCCGATGCTAGCGCTCTGCGCTTGGGCTCTACGGG 215
Qy 181 GCGCTTCTTTTACGCGACCTGGTGGCGAGAGCTCTTTCGCTACCTTGAGCAGCGGG 240
Db 216 GCGCTTCTTTTACGCGACCTGGTGGCGAGAGCTCTTTCGCTACCTTGAGCAGCGGG 275
Qy 241 GTGGCGGGCGGGCGGGCGGGCGCTGATGAGCGACCGCGCGAGTGTGGCGCTGACC 300
Db 276 GTGGCGGGCGGGCGGGCGGGCGCTGATGAGCGACCGCGCGAGTGTGGCGCTGACC 335
Qy 301 ATCTCGGCTTACGAGAGAGCGCGGCTGCTGCTGCGGCTGCTGCGCTGCGCGCGCC 360
Db 336 ATCTCGGCTTACGAGAGAGCGCGGCTGCTGCTGCGGCTGCTGCGCTGCGCGCGCC 395
Qy 361 CTGCTGTATACCGCGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 396 CTGCTGTATACCGCGCGCGGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 455
Qy 421 GACCTCTACATGCTGCAATGCTTCCGAGAGTCTTCTGCTGAGAGAGAGCGCGCGTAC 480
Db 456 GACCTCTACATGCTGCAATGCTTCCGAGAGTCTTCTGCTGAGAGAGAGCGCGCGTAC 515
Qy 481 GTGTGGAGCGGCACTACCACTGAGAGAGAGCGCGCGGCGGGCGGGCGGGCGGG 540
Db 516 GTGTGGAGCGGCACTACCACTGAGAGAGAGCGCGCGGCGGGCGGGCGGGCGGG 575
Qy 541 GGAGGCTATCGGAGGCTGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGCGCTGTGTG 600
Db 576 GGAGGCTATCGGAGGCTGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGCGCTGTGTG 635
Qy 601 AGGACTCGAGGCTGT 660
Db 636 AGGACTCGAGGCTGT 695
Qy 661 ACAGCTTCAAGGCGCTCGAGATTCGGTGAGTCTGCTGAGTCTGCTGAGTCTGAGTCTG 720
Db 696 ACAGCTTCAAGGCGCTCGAGATTCGGTGAGTCTGCTGAGTCTGCTGAGTCTGAGTCTG 755
Qy 721 AGGTTGGACCCCATGGCACTGCTGAGCTCTGCGGCTGCTGAGTCTGAGTCTGAGTCTG 780
Db 756 AGGTTGGACCCCATGGCACTGCTGAGCTCTGCGGCTGCTGAGTCTGAGTCTGAGTCTG 815
Qy 781 GGGGCTGT 840
Db 816 GGGGCTGT 875
Qy 841 AGCAGCTGCGATCTGAGTGTGCTTCAATGTGGAGCGGCTTGTGAGTCTGAGTCTTCCAC 900
Db 876 AGCAGCTGCGATCTGAGTGTGCTTCAATGTGGAGCGGCTTGTGAGTCTGAGTCTTCCAC 935
Qy 901 TGTGTATCTGCTCAGCGGCTCTAG----- 928
Db 936 TGTGTATCTGCTCAGCGGCTCTAG----- 995
Qy 929 -----AATCTGCCAGGCGCCAGGGAGCAGCGATGATGCC 965
Db 996 CTTGAGGCTGTGTAACAACAGAGTCTCTGGGTACCCACTGACTTTTGGGGATGACCGG 1055
Qy 966 CTCATCTCTCGCCC-----CGTCCAGGTACACTCTCAGGTCGCGC 1006
Db 1056 CACCTCAACAACCGCATGCTCAGCATGGGTATGTCTCAAGATACACTCTCAGGTCGCGC 1115
Qy 1007 TGCTACTCAGAGACCGCTCTGCTTCTGCGGCTGGCTGAGCAGCAGCAGCTGTGTCC 1066
Db 1116 TGCTACTCAGAGACCGCTCTGCTTCTGCGGCTGGCTGAGCAGCAGCAGCTGTGTCC 1175

Qy 1067 AAGTCGTACTTCCGTGA 1083
Db 1176 AAGTCGTACTTCCGTGA 1192
RESULT 2
US-09-949-016-5659
; Sequence 5659, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5659
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5659

Query Match 85.9%; Score 929.8; DB 4; Length 2087;
Best Local Similarity 90.4%; Pred. No. 4.1e-175;
Matches 1046; Conservative 0; Mismatches 37; Indels 74; Gaps 2;
Qy 1 ATGAGACGACGAGCGCGCCCAAGCCCACTCTCTGACGCGCGGCTGCTCGGCGCTGACC 60
Db 36 ATGAGACGACGAGCGCGCGCCCAAGCCCACTCTCTGACGCGCGGCTGCTCGGCGCTGACC 95
Qy 61 CGAGGGTGTGACCATGCGCTTCCGCTGCTCATCTTGGGCTCATGACCTGGGCTTAC 120
Db 96 CGAGGGTGTGACCATGCGCTTCCGCTGCTCATCTTGGGCTCATGACCTGGGCTTAC 155
Qy 121 GCGCGCGGGTGGCGCTGCGCTCCGATGCTAGCGCTCTGCGCTTGGGCTCTACGGG 180
Db 156 GCGCGCGGGTGGCGCTGCGCTCCGATGCTAGCGCTCTGCGCTTGGGCTCTACGGG 215
Qy 181 GCGCTTCTTTTACGCGACCTGGTGGCGAGAGCTCTTTCGCTACCTTGAGCAGCGGG 240
Db 216 GCGCTTCTTTTACGCGACCTGGTGGCGAGAGCTCTTTCGCTACCTTGAGCAGCGGG 275
Qy 241 GTGGCGGGCGGGCGGGCGGGCGCTGATGAGCGACCGCGCGAGTGTGGCGCTGACC 300
Db 276 GTGGCGGGCGGGCGGGCGGGCGCTGATGAGCGACCGCGCGAGTGTGGCGCTGACC 335
Qy 301 ATCTCGGCTTACGAGAGAGCGCGGCTGCTGCTGAGTCTGCGGCTGCGCGCGCGCC 360
Db 336 ATCTCGGCTTACGAGAGAGCGCGGCTGCTGCTGAGTCTGCGGCTGCGCGCGCGCC 395
Qy 361 CTGCTGTATACCGCGCGGCTGCGGCTCTCATGTGTGGATGGCAACCGCGCGGAG 420
Db 396 CTGCTGTATACCGCGCGGCTGCGGCTCTCATGTGTGGATGGCAACCGCGCGGAG 455
Qy 421 GACCTCTACATGCTGCAATGCTTCCGAGAGTCTTCTGCTGAGAGAGAGCGCGCGTAC 480
Db 456 GACCTCTACATGCTGCAATGCTTCCGAGAGTCTTCTGCTGAGAGAGAGCGCGCGTAC 515
Qy 481 GTGTGGAGCGGCACTACCACTGAGAGAGAGCGCGCGGCGGGCGGGCGGGCGGG 540
Db 516 GTGTGGAGCGGCACTACCACTGAGAGAGAGCGCGCGGCGGGCGGGCGGGCGGG 575
Qy 541 GGAGGCTATCGGAGGCTGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGCGCTGTGTG 600

Db 576 GGAGCCTATCGGAGGTGGAGCGGAGGATTCCTGGGCGCTGGACGTGAGGCGCTGGTG 635
Qy 601 AGGACTCGCAGGTGCTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGCAGGTTCATGTAC 660
Db 636 AGGACTCGCAGGTGCTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGCAGGTTCATGTAC 695
Qy 661 ACAGCCTTCAAGCGCTCGGAGATTGCTGGTGAGTACTGTCGAGGTCTGTGACTCGGACACA 720
Db 696 ACAGCCTTCAAGCGCTCGGAGATTGCTGGTGAGTACTGTCGAGGTCTGTGACTCGGACACA 755
Qy 721 AGTTGGACCCCATGCACTGCTGGAGCTGCTGGGCTACTGAGCAGGACCCCGGGTA 780
Db 756 AGTTGGACCCCATGCACTGCTGGAGCTGCTGGGCTACTGAGCAGGACCCCGGGTA 815
Qy 781 GGGGCTGTTGGTGGGACGTGGGATTCCTTAACCCCTCTGGACTCTCTGGGTCAAGTTCCTA 840
Db 816 GGGGCTGTTGGTGGGATGTCGGATCCTTAACCCCTCTGGACTCTCTGGGTCAAGTTCCTA 875
Qy 841 AGCAGCTCGGATCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCAC 900
Db 876 AGCAGCTCGGATCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCAC 935
Qy 901 TGTGTATCTGTCATCAGCGGTCTCTAG----- 928
Db 936 TGTGTATCTGTCATCAGCGGTCTCTAGGGCTATATAGGAATAACCTCTTGCAGCAGTTT 995
Qy 929 -----AATCCTGCCAGCGCCCGGAGCAGCGCATGATGCC 965
Db 996 CTTGAGGCTGTGACAAACAGAGTTCCTGGGTACCCACTGTACTTTTGGGATGACCGG 1055
Qy 966 CTCATTCTCGCCCC-----CGTGAGGTACACCTCCAGGTCCCGC 1006
Db 1056 CACCTCACCAACCGCATGCTCAGCATGGTATTATGTACCAAGTACACCTCCAGGTCCCGC 1115
Qy 1007 TGCTACTCAGACCGCTCTGCTCTGCTGGTGGCTGAGCAGCAGACAGCTGGTCC 1066
Db 1116 TGCTACTCAGACCGCTCTGCTCTGCTGGTGGCTGAGCAGCAGACAGCTGGTCC 1175
Qy 1067 AAGTCGTACTTCCGTGA 1083
Db 1176 AAGTCGTACTTCCGTGA 1192

RESULT 3

US-09-155-768-1
; Sequence 1, Application US/09155768A
; Patent No. 6162908
; GENERAL INFORMATION:
; APPLICANT: SEIKAGAKU CORPORATION
; TITLE OF INVENTION: POLYPEPTIDE OF HYALURONAN SYNTHASE OF
; TITLE OF INVENTION: HUMAN-ORIGIN AND DNA ENCODING THE SAME
; FILE REFERENCE: TOVAM37.001APC
; CURRENT APPLICATION NUMBER: US/09/155,768A
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: JP 8-084326
; EARLIER FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: JP 8-109663
; EARLIER FILING DATE: 1996-04-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2117
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (149)...(1777)
US-09-155-768-1

Query Match 85.0%; Score 920.6; DB 3; Length 2117;
Best Local Similarity 90.2%; Pred. No. 2.7e-173;
Matches 1038; Conservative 0; Mismatches 39; Indels 74; Gaps 2;

Qy 7 CAGCAGGACGCGCCCAAGCCCACTCCTGACGAGCCCGCCGCTGCTCGGCGCTGCGCCGAGG 66
Db 50 CAGCAGGACGCGCCCAAGCCCACTCCTGACGAGCCCGCCGCTGCTCGGCGCTGCGCCGAGG 109
Qy 67 GTGCTGACCATCGCCTTGGCCTGTCTATCTCGGGGCTCATGACCTGGGCTTACGCGGC 126
Db 110 GTGCTGACCATCGCCTTGGCCTGTCTATCTCGGGGCTCATGACCTGGGCTTACGCGGC 169
Qy 127 GGGGTGCGGCTGCGCTTCGATCGCTACGGGCTCTCGGCTTCTCGGCTTCTACGGGGCTTTC 186
Db 170 GGGGTGCGGCTGCGCTTCGATCGCTACGGGCTCTCGGCTTCTCGGCTTCTACGGGGCTTTC 229
Qy 187 CTTTCAGCGCACCTGGTGGCGCAGAGCTCTTTCGCTACTCTGGAGACCGCGGGGTGGG 246
Db 230 CTTTCAGCGCACCTGGTGGCGCAGAGCTCTTTCGCTACTCTGGAGACCGCGGGGTGGG 289
Qy 247 GCGGGCGCGCGGGGCGCTGGATGACGCCACCGCGCGCAGTGTGGCGCTGACCATCTCC 306
Db 290 GCGGGCGCGCGGGGCGCTGGATGACGCCACCGCGCGCAGTGTGGCGCTGACCATCTCC 349
Qy 307 GCCTACACGAGGACCCCGCGTACCTGCGCCAGTCCCTGGGCTCCCGCGCCCTGCTG 366
Db 350 GCCTACACGAGGACCCCGCGTACCTGCGCCAGTCCCTGGGCTCCCGCGCCCTGCTG 409
Qy 367 TACCGCGCGCGCGCTCGGCTCCTCATGTGTGGATGCAACCGCGCGCAGAGACTC 426
Db 410 TACCGCGCGCGCGCTCGGCTCCTCATGTGTGGATGCAACCGCGCGCAGAGACTC 469
Qy 427 TACATGCTGACATGTTTCGGGAGGTCTTCGCTGACGAGGACCCCGCGCAGTGTGG 486
Db 470 TACATGCTGACATGTTTCGGGAGGTCTTCGCTGACGAGGACCCCGCGCAGTGTGG 529
Qy 487 GACGCAACTACACACGCTTGGAAACCGCGCGCGCGGCTGGGCGCCGAGGC 546
Db 530 GACGCAACTACACACGCTTGGAAACCGCGCGCGCGGCTGGGCGCCGAGGC 589
Qy 547 TATCGGAGGTGGAGCGCGAGATCCTGGGCGGCTGGCAGTGGAGCGCTGTGAGGACT 606
Db 590 TATCGGAGGTGGAGCGCGAGATCCTGGGCGGCTGGCAGTGGAGCGCTGTGAGGACT 649
Qy 607 GCGAGGTGCTGCTGGCGCGCGCTGGGGCGCAAGCGCGAGGTCTATGACAGCC 666
Db 650 GCGAGGTGCTGCTGGCGCGCGCTGGGGCGCAAGCGCGAGGTCTATGACAGCC 709
Qy 667 TTCAAGGCGCTCGGAGATTTCGCTGACATCTGACAGGTCTGTGACTCGACACAAGTTG 726
Db 710 TTCAAGGCGCTCGGAGATTTCGCTGACATCTGCTGAGGTCTGTGACTCGACACAAGTTG 769
Qy 727 GACCCATGGCAGTCTGCGAGCTCGTGGGCTACTGGAGGACCCCGGGTAGGGCT 786
Db 770 GACCCATGGCAGTCTGCTGGGCTACTGGAGGACCCCGGGTAGGGCT 829
Qy 787 GTTGGTGGGAGGTGCGGATCCTTAACCCCTGGACTCTGCTGGTCTGCTTCTTAAGCAGC 846
Db 830 GTTGGTGGGAGTGTGCGGATCCTTAACCCCTGGACTCTGCTGGTCTGCTTCTTAAGCAGC 889
Qy 847 CTGCGATCTGCTGAGCTTCAATGTGGAGCGGCTTGTGACAGCTACTTCCACTGTGA 906
Db 890 CTGCGATCTGCTGAGCTTCAATGTGGAGCGGCTTGTGACAGCTACTTCCACTGTGA 949
Qy 907 TCCTGCTCAGCGGTCTCTAG----- 928
Db 950 TCCTGCTCAGCGGTCTCTAGGGCTATATAGGAATAACCTTTCGACAGTCTTCTTGAG 1009
Qy 929 -----AATCCTGCCAGGCGCCCGGAGCAGCGGATGATGCGCTCATTT 971
Db 1010 GCTGTGTAACAAGAGGTTCCTGGGTACCCACTGTACTTTTGGGGATGACCGGCACTC 1069
Qy 972 CCGCGCCC-----CGTGAGGTACCTCCAGGTCCCGCTGCTAC 1012
Db 1070 ACCAAGCGATGCTCAGCATGGGTATGCTACCAAGTACCTCCAGGTCCCGCTGCTAC 1129
Qy 1013 TCAGAGCGCCTCGTCTCTCTGCGGTGGCTGAGCCAGCAGACACGCTGGTCCCAAGTCG 1072

Db 1130 TCAGAGACGCCCTCGTCTTCTCGCGTGGCTGAGCCAGCAGACACGCTGGTCCAAGTCG 1189

Qy 1073 TACTTCGTGA 1083

Db 1190 TACTTCGTGA 1200

RESULT 4

```

US-09-949-016-17401
; Sequence 17401, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17401
; LENGTH: 14862
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17401

```

Qy	547	TATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGTGAGCGCTGTGTGAGGACT	606
Db	6613	TATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGTGAGCGCTGTGTGAGGACT	6672
Qy	607	CGCAGTGTGCTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGCAGTCTATGTTCACAGCC	666
Db	6673	CGCAGTGTGCTGTGCTGGCGCAGCGCTGGGGCGGCAAGCGCAGTCTATGTTCACAGCC	6732
Qy	667	TTCAAGCGCGCTCGGAGATTCGGTGGAGTACGTGTCAGGTCCTGT	708
Db	6733	TTCAAGCGCGCTCGGAGATTCGGTGGAGTACGTGTCAGGTCCTGT	6774

5 T.M.S.A

```

RESUL 3
US-09-949-016-201775/c
; Sequence 201775, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 201775
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-201775

```

RESULT 6

US-09-949-016-3728

Db	652	ACTGTGCTGGATCCAGCGCTTGCA	CAATCGAGATGCTTCGAGTCTCTGGAGAGGAGATCCCCAA	711
Qy	778	GTAGGGCGCTGTTGGTGGGGACGTCGCGGATCCTTAAACCTCTCTGGACTCTCTGGGTGAGCTTC	837	
Db	712	GTAGGGGGAGTGGGGGAGATGCCAGATCCTCAACAAGTACGACTCATGGATTTCCCTTC	771	
Qy	838	CTAAGCAGCGCTCGGATACTGGGGTAGCCTTCAAATGTGGAGGGGGCTTGTTCAGAGCTACTTC	897	
Db	772	CTGAGCAGCGTGGTACTGATGGCTTCAACGCTGGAGGGGCTGCGAGTCTCTACTTT	831	
Qy	898	CACGTGTATCTCTGCATCAGCGGTCTCTTAG	928	
Db	832	GGCTGTGCGAGTGTATTAGTGGCCCTTGG	862	
RESULT 7				
US-08-812-008-31				
; Sequence 31, Application US/08812008				
; Patent No. 6602693				
; GENERAL INFORMATION:				
; APPLICANT: McDonald, J. A.				
; APPLICANT: Spicer, A. P.				
; APPLICANT: Augustine, M. L.				
; TITLE OF INVENTION: GENE ENCODING HYALURONAN				
; TITLE OF INVENTION: SYNTHASE				
; NUMBER OF SEQUENCES: 54				
; CORRESPONDENCE ADDRESS:				
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.				
; STREET: P.O. Box 2938				
; CITY: Minneapolis				
; STATE: MN				
; COUNTRY: U.S.A				
; ZIP: 55402				
; COMPUTER READABLE FORM:				
; MEDIUM TYPE: Diskette				
; COMPUTER: IBM Compatible				
; OPERATING SYSTEM: DOS				
; SOFTWARE: FastSEQ Version 2.0				
; CURRENT APPLICATION DATA:				
; APPLICATION NUMBER: US/08/812,008				
; FILING DATE: 05-MAR-1997				
; CLASSIFICATION:				
; PRIOR APPLICATION DATA:				
; APPLICATION NUMBER: 08/675,499				
; FILING DATE: 03-JUL-1996				
; ATTORNEY/AGENT INFORMATION:				
; NAME: Embretson, Janet E				
; REGISTRATION NUMBER: 39,665				
; REFERENCE/DOCKET NUMBER: 150.183US1				
; TELECOMMUNICATION INFORMATION:				
; TELEPHONE: 612-373-6900				
; TELEFAX: 612-339-3061				
; TELEX:				
; INFORMATION FOR SEQ ID NO: 31:				
; SEQUENCE CHARACTERISTICS:				
; LENGTH: 1665 base pairs				
; TYPE: nucleic acid				
; STRANDEDNESS: unknown				
; TOPOLOGY: unknown				
; MOLECULE TYPE: cDNA				
; FEATURE:				
; NAME/KEY: Coding Sequence				
; LOCATION: 1..1662				
; OTHER INFORMATION:				
; US-08-812-008-31				
Query Match 21.8%; Score 236.4; DB 4; Length 1665;				
Best Local Similarity 58.5%; Pred. No. 5.7e-38;				
Matches 511; Conservative 0; Mismatches 321; Indels 42; Gaps 4;				
Qy	61	CGAGGGGTGTCACCATCGCGCTTCGCGCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC	120	
Db	28	CGTGTGGTGGGACCAAGTCTGTTGCGCTGGTAGTGTCTGGAGGACATCCTGGCGGCTAT	87	

Qy 121 GCCGCGGGTGGCGCTGCCATCGCTACGCGCTCTGGCTCTGGCTCTACGGG 180
Db 88 GTGACAGGCTACCAAGTTTATCCACAGAAAGACATACCTGTCTCTTGGCTCTACCGT 147
Qy 181 GCCTTCTCTTTTCAGCGCACCTGTGGCGCAGAGCTCTTCGCGTACCTGGAGCACCGCGG 240
Db 148 GCCATCTCTGGTCTACATCTGCTCATCCAGAGCTGTCTTCTCTGGAGCACCGTCA 207
Qy 241 GTGGGGGGGGGGCGG-----GGGGCGCTGATGAGCCACCGCGCGAGTGTGGCG 294
Db 208 ATGGCGAGGCGAGCGCGCGCCCTCAAGCTGCACTGCTCCAGAGAGTGCCTTCAGTGGCA 267
Qy 295 CTGACCATCTCCGCTACACAGAGAGCCCGCTACCTGCGCCAGTGCCTGGCTCGGCC 354
Db 268 CTCTGCAATGCTGCTTACCAAGAGACCCCGAAATACCTGCGCAAGTGCCTTCGCTCAGCT 327
Qy 355 CGCGCCCTGTGTACCCGCGCGCGGCTGCGCGCTCTCATGTGTGGATGGCAACCGC 414
Db 328 CAGCGCATGCTCTTCCAAAC-----CTCAAGGTGGTCAATGATGGTGGCAATCGC 381
Qy 415 GCCAGGACCTCTACATGGTTCGATGTTCCGCGAGGTCTTCGCTGAGAGGACCCCGCC 474
Db 382 CAGGAAGATACCTACATGTTGGACATCTTCCATGAGGTGCTGGGTGGCACTGAGCAAGCT 441
Qy 475 ACGTACGTGTGGAGCGCACTACACAGCCCTGGGNAACCGCGGGGGCGCGGTG 534
Db 442 GGCTCTTGTGTGGGTGAGCAATTTCCAT-----GAGGCGGGTGAAGGA 486
Qy 535 GCGCGGAGCTTATCGGAGGTGAGGCGGAGGATCTTGGCGGCTGGCAGTGGAGCG 594
Db 487 GAGACAGGCGCGCTGCAGAGGCAATGAGCGT-----GTGCGAGCT 531
Qy 595 CTGTGAGGACTCGAGGTGCTGTGTGCGCGCAGCGCTGGCGGCGCAAGCGCAGGTC 654
Db 532 GTGGTGTGGCGCAGCACCTTCTCATGCAATCATGCAAGATGGGGGGCGCAAGCGTGAAGTC 591
Qy 655 ATGTACACAGCTTCAGGCGCTCGGAGATTCGGTGGCTGACGTGCGAGTCTGTGACTCG 714
Db 592 ATGTACACTGCTTCAGGCGCTTGGCACTTCAGTGACATACCTCAGAGTGTGACTCT 651
Qy 715 GACACAGGTGTGGACCCCATGCGACTGTGTGAGCTGTGTGGGTGCTGAGAGGACCCC 774
Db 652 GACACTGTGTGGACCCCATGCGACTTCAGATGCTTCGAGTCTTGGAGAGATCCC 711
Qy 775 CGGTAGGGCTGTGTGGGAGCTGGGATCTTAACTCTGAGATCTCTGGGTGAGC 834
Db 712 CAAGTAGGAGGTGTGGAGGAGATGTCCAAATCCTCAACAGTATGATTCATGATCTCC 771
Qy 835 TTCTTAAGCAGCTCGCATCTGGTGTAGCCCTTCAATGTGGAGCGGCTTGTGAGGATC 894
Db 772 TTCTGAGCAGTGTGAGTGTGATGCTTCAACGTGGAGCGGCGCTGCCAGTCTTAC 831
Qy 895 TTCCACTGTGTATCTCTGCATCAGCGGCTCTCTAG 928
Db 832 TTTGGCTGTGCAATGATTAAGTGGGCTTTGG 865

RESULT 8

US-09-949-016-201780/c
; Sequence 201780, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201780
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-201780

Query Match 21.2%; Score 229.4; DB 4; Length 601;
Best Local Similarity 98.7%; Pred. No. 1.3e-36;
Matches 230; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 696 CGTCAGGTCGTGACTCGGACACAAAGTTGGACCCCATGGCACTGCTGGAGCTCGTGGC 755
Db 403 CTTACAGGTCGTGACTCGGACACAAAGTTGGACCCCATGGCACTGCTGGAGCTCGTGGC 344
Qy 756 GGTACTGACGAGGACCCCGGGTAGGGGTGTGGTGGGACGCTGGCGATCCTTAACCC 815
Db 343 GGTACTGACGAGGACCCCGGGTAGGGGTGTGGTGGGAYGTGGATCCTTAACCC 284
Qy 816 TCTGAGCTCTGGGTGAGCTTCCTAAGCAGCTCGCATCTGGGTAGCCTTCAATGTGA 875
Db 283 TCTGAGCTCTGGGTGAGCTTCCTAAGCAGCTCGCATCTGGGTAGCCTTCAATGTGA 224
Qy 876 GCGGGCTTGTGAGGCTACTTCCACTGTGTATCTGTCATCAGCGGTCTCTAG 928
Db 223 GCGGGCTTGTGAGGCTACTTCCACTGTGTATCTGTCATCAGCGGTCTCTAG 171

RESULT 9

US-08-675-499A-1
; Sequence 1, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwesman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2947 base pairs
; TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-675-499A-1

Query Match 16.6%; Score 179.4; DB 4; Length 2947;
Best Local Similarity 55.3%; Pred. No. 1.1e-26;
Matches 441; Conservative 0; Mismatches 311; Indels 45; Gaps 3;

Qy 159 CTGGCTTGGGCTTACGGGGCTTCTTTCAGCGCACCTGGTGGCGCAGAGCTCTT 218
Db |||||
Qy 636 CTTCTCATTTGGACTGAGTGCTCTTTTACCTCGCATCTCATCTCAAGAGCTCTT 695
Db |||||
Qy 219 CGCGTACTGGAGCACCGCGGGTGGCGCGCGCGCGCGCGCGCTGGATGCAGCCAC 278
Db |||||
Qy 696 TGCCTTTTGGAAACACCGGAAATGAAGAAGTCCCTTGAACCCCGATTAAATGAACAA 755
Db |||||
Qy 279 CGCGGAGTGTGGGCTGACATCTCCGCTACAGGAGACCCCGGCTACCTGGCGCA 338
Db |||||
Qy 756 AACG-----GTAGCACTCTGCATCGCTGCGTACCAAGAGGACCTGACTACTACGGA 809
Db |||||
Qy 339 GTGCTGCGCTCGCGCGCGCTGCTGTACCGCGCGCGCGCGCGCTGCGCTCCTCATGCT 398
Db |||||
Qy 810 ATGTTTGCATCTGTGAAGAGCTGACCTACCTCG-----GGATTAAAGTCGTATGCT 863
Db |||||
Qy 399 GGTGGATGGCAACCGCGCGCGAGGACCTCTCATGCTGACATGTTCCGCGAGGCTTTCGC 458
Db |||||
Qy 864 CATCGATGGGAACTCAGACGACGACCTTTACATGATGACATATTCACGGAATATTGG 923
Db |||||
Qy 459 TGACGAGACCCCGCCACGCTACGCTGTGGGAGCGGCAACTACACAGCCCTGGGAACCCGC 518
Db |||||
Qy 924 CAGGACAAATCGGCGACGATCATCTGGAAGAACTTTCATGAAAAG----- 972
Db |||||
Qy 519 GCGCGCGCGCGCTGGGCGCGCGAGCTATCGGAGGTGGAGGCGGAGGATCTTGGCG 578
Db |||||
Qy 973 -----GGACCTGGTGAGACAGAGATCCCATATAAGAAAGTTC 1010
Db |||||
Qy 579 GCTGGCAGTGGAGGCGCTGGTGGAGTCTCGCAGGTGGTGTGCGTGGCGCAGCGCTGGG 638
Db |||||
Qy 1011 ACAACATGTACCCCAATTTGGTCTTGTCTAAACAAAGTATTTGCATCATGCAAAATGGG 1070
Db |||||
Qy 639 CGCAAGCGCAGGTCATGTACACAGCTTCAAGCGCTTCGAGATTCGGTGGAGTACGT 698
Db |||||
Qy 1071 TGAAGAGAGAAAGTCATGTACACAGCTTCAGAGCACTGGGCGGAAAGCGTGGATTATGT 1130
Db |||||
Qy 699 GCAGTCTGACCTGGACACAGTTCGACCCCATGCTGCTGGAGCTCGCGGT 758
Db |||||
Qy 1131 ACAGGTGTGACTCAGATATATGCTTGAACCTGCTGCTGCTGAGAGTGGTGAAGGT 1190
Db |||||
Qy 759 ACTGACGAGACCCCGGCTAGGGCTGTGGTGGGACGCTGGGATCTTAAACCTCT 818
Db |||||
Qy 1191 CTTAGAGGAAGCCCTATGTTGGAGGTGTTGGAGGAGATGTCAGATTTTAAACAGTA 1250
Db |||||
Qy 819 GACCTCTGGGTGAGTTCCTAAGCAGCTCGATCTAGGCTGAGTTCATGAGTGGAGCG 878
Db |||||
Qy 1251 TGAATCTGGATCTCTTCTCAGCAGCGTGAGATCTGATGCTTTTAAATATAGAAAG 1310
Db |||||
Qy 879 GGCTGTGACAGTACTTCCACTGTATCTGCAACAGCGTCTCTAGATCTTGGCC 938
Db |||||
Qy 1311 GGCCTGCGAGTCTTATTTTGGCTGTGCTCAGTGCAAGAGCGCTCTTGGGAATGTACAG 1370
Db |||||
Qy 939 AGGCCCGAGGAGACG 955
Db |||||
Qy 1371 AAATCTTGTGCAATG 1387
Db |||||

RESULT 10

US-08-812-008-1

; Sequence 1, Application US/08812008
; Patent No. 6602693
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.

APPLICANT: Augustine, M. L.
TITLE OF INVENTION: GENE ENCODING HYALURONAN
TITLE OF INVENTION: SYNTHASE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwesman, Lundberg, Woessner & Kluth, P.A.
STREET: P.O. Box 2938
CITY: Minneapolis
STATE: MN
COUNTRY: U.S.A.
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/812,008
FILING DATE: 05-MAR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/675,499
FILING DATE: 03-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Embretson, Janet E
REGISTRATION NUMBER: 39,665
REFERENCE/DOCKET NUMBER: 150.183US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-373-6900
TELEFAX: 612-339-3061
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2947 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-812-008-1

Query Match 16.6%; Score 179.4; DB 4; Length 2947;

Best Local Similarity 55.3%; Pred. No. 1.1e-26;

Matches 441; Conservative 0; Mismatches 311; Indels 45; Gaps 3;

Qy 159 CTGGCTTGGGCTTACGGGGCTTCTTTCAGCGCACCTGGTGGCGCAGAGCTCTT 218
Db |||||
Qy 636 CTTCTCATTTGGACTGAGTGCTCTTTTACCTCGCATCTCATCTCAAGAGCTCTT 695
Db |||||
Qy 219 CGCGTACTGGAGCACCGCGGGTGGCGCGCGCGCGCGCGCTGGATGCAGCCAC 278
Db |||||
Qy 696 TGCCTTTTGGAAACACCGGAAATGAAGAAGTCCCTTGAACCCCGATTAAATGAACAA 755
Db |||||
Qy 279 CGCGGAGTGTGGGCTGACATCTCCGCTACAGGAGACCCCGGCTACCTGGCGCA 338
Db |||||
Qy 756 AACG-----GTAGCACTCTGCATCGCTGCGTACCAAGAGGACCTGACTACTACGGA 809
Db |||||
Qy 339 GTGCTGCGCTCGCGCGCGCTGCTGTACCGCGCGCGCGCGCTGCGCTCCTCATGCT 398
Db |||||
Qy 810 ATGTTTGCATCTGTGAAGAGCTGACCTACCTCG-----GGATTAAAGTCGTATGCT 863
Db |||||
Qy 399 GGTGGATGGCAACCGCGCGCGAGGACCTCTCATGCTGACATGTTCCGCGAGGCTTTCGC 458
Db |||||
Qy 864 CATCGATGGGAACTCAGACGACGACCTTTACATGATGACATATTCAGGGAATATTGG 923
Db |||||
Qy 459 TGACGAGACCCCGCCACGCTACGCTGTGGGAGCGGCAACTACACAGCCCTGGGAACCCGC 518
Db |||||
Qy 924 CAGGACAAATCGGCGACGATCATCTGGAAGAACTTTCATGAAAAG----- 972
Db |||||
Qy 519 GCGCGCGCGCGCTGGGCGCGCGAGCTATCGGAGGTGGAGGCGGAGGATCTTGGCG 578
Db |||||
Qy 973 -----GGACCTGGTGAGACAGAGATCCCATATAAGAAAGTTC 1010
Db |||||
Qy 579 GCTGGCAGTGGAGGCGCTGGTGGAGTCTCGCAGGTGGTGTGCGTGGCGCAGCGCTGGG 638
Db |||||

1011	ACAACATGTACCCAAATTGGTCTTGTCTAAACAAAGTATTTCATCATGTCAAAAATGGG	1070
Qy	CGGCAAGCGGAGTGCATGTACACAGCCTTCAAGCGCTCGGAGATTTCGGTGGACTACGT	698
Db	TGGAAAGAGAGAAGTCAATGTACACAGCCTTCAGAGCACTGGGGCGAAGCGTGGATTATGT	1130
Qy	GCAGTCTGTGACTCGGACACAAAGTTGGACCCCATGGCACTGTCGAGACTCGTGC GGGT	758
Db	ACAGTGTGTACTCAGATACATATGCTTGACCCCTGCCTCATCTGTGGAGATGGTGAAGGT	1190
Qy	ACTGCAGCAGGACCCCGGTAGGGGCTGTTGGTGGGACGTGGCGATCCTTAAACCCCTCT	818
Db	CTTAGAGGAAGACCTATGTTGGAGGTGTTGGAGGAGATGTCAGATTTTAAACAAGTA	1250
Qy	GGACTCTCGGTTCAGTCTCTTAAGCAGCTCGATACTGGGTAGCCTTCAATGTGGAGCG	878
Db	TGATTCTTGGATCTCCTTCTCAGCAGGTGAGATACTGGATGGCTTTTAATATAGAAG	1310
Qy	GGCTTGTACAGACTACTTCCACTGTGATCCTGCATCAGCGGTCTCTTAGAATCCTGCC	938
Db	GGCCTGCCAGTCTTATTTTGGCTGTCTCCAGTGCATAAGCGGTCTCTGGGAATGTACAG	1370
Qy	AGGCCCCAGGAGCAGC	955
Db	AAACTCTTGTCTGCATG	1387

```

RESULT 11
US-09-949-016-15470
; Sequence 15470, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15470
; LENGTH: 9871
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15470

```

Qy	298	ACCAATCCGGCTACAGAGAGAGACCCCGCGTACTTGGCGCAGTGGCTTGGCGGTCCGCCCGC	357
Db	2368	TGCAATCCGCGATACCAAGAGAGACCCCTGACTTACTTGGCGAAGTGGCTTGGCTCGGCCCCAG	2327
Qy	358	GCCTGTGTATCCGCGCGCGCGCTCGCGTCTCTCATGTGTGTGGATGCAACCGCGCC	417
Db	2328	CGCATCTCTTCCTCCCTGAC-----CTCAAGGTGGTCAATGTGTGGATGGCAACCGCAG	2381
Qy	418	GAGGACCTCTACATGTGTGCATGTTCCGCGAGGTCTTCGTGACAGGAGACCCGCCACG	477
Db	2382	GAGGACGCTTACATGCTGGACATCTTCCAGAGGTGTGGGCGGCACCGAGAGCGCCGC	2441
Qy	478	TACGTGTGGAGGCGCAACTACACGCCCTTGGGAACCCGCGCGCGCGCGTGGCG	537
Db	2442	TTCCTTTGTGTGGCGCAGCAACTTCCAT-----GAGGCAGGCGAGGGTCAG	2486
Qy	538	GCCGGAACCTATCGGGAAGGTGGAGCGGAGGATCCTGGCGGTGGCAGTGGAGGCGCTG	597
Db	2487	ACGAGAGCCACGCTTGCAGGAGGGCATGGACCGT-----GTGCGGATGTG	2531
Qy	598	GTGAGGACTCGCAGGTCCGTGTGGTGGCGCAGCGCTGGGCGCGCAAGCGCAGGTCATG	657
Db	2532	GTGCGGCGCAGACCTTCTCGTGCATCATGCAGAAAGTGGGAGGCAAGCGCAGGTCAATG	2591
Qy	658	TACACAGCCTTCAAGGCGCTCGGAGATTCGGTGGACTACTGTGACGTTCTGTG	709
Db	2592	TACAGGCTTCAAGGCCCTCGGCGATTCGGTGGACTACATCCAGGTAAAGG	2643

RESULT 12

```

RES001 12
US-09-949-016-737
; Sequence 737, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 737
; LENGTH: 3003
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-737

```

Qy	404	ATGCAACCGCGCCGAGGACTCTACATGTGTGACATGTTCCGGGAGGCTTTCGCTGACG	463
Db	897	ATGGGAATCTCAGAAGATGACCTTTACATGATGGACATCTTCAGTGAAGTCATGGGCAGAG	956
Qy	464	AGGACCCGCGCAGCTACGTGTGTGGGACGGCAACTACCAACGACCCCTGGGACCCCGCGCGG	523
Db	957	ACAATCAGCCACTTATATCTGGAGAGAACAACTTCCACGAAAG-----	1000
Qy	524	CGGCGCGGTGGGCGCGGAGCCCTATCGGAGGTGGAGGGGAGGATCTCTGGCGGCTGG	583
Db	1001	-----GGTCCCGGTGAGACAGATGAGTCACATAAGAAGCTCGCAAC	1043
Qy	584	CAGTGGAGGCGCTGGTGAGGACTCGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGCGGCGCA	643
Db	1044	ACGTACGCAATTTGGTCTGTGTCACAAACAAAGTATCTGCATCATGCAAAAATGGGTGGAA	1103
Qy	644	AGCGGAGGTCAATGTAACAGACCTTCAAGGGCGTCGGAGATTGCGTGGGACTACGTGCAAG	703
Db	1104	AAAGAGAAGTCATGTACACAGCCTTCAGAGCACTGGGACGAAGTGTGGATTATGTACAGG	1163
Qy	704	TCTGTGACTCGGACACAAGTTGGACCCCATGGCACTCTGTGGAGCTCGTGGGGTACTGG	763
Db	1164	TTTGTGATTACAGACACTATGCTTGACCAGCCTCATCTGTGGAGATGGTAAAAAGTTTTAG	1223
Qy	764	ACGAGGACCCCGGTAGGGGCTGTTGGTGGGAGCGTGGGATCCTTAAACCTCTGGACT	823
Db	1224	AAGAAGATCCCATGGTTGGAGGTTGGGGGAGATGTCAGATTTTAAACAAGTACGATT	1283
Qy	824	CTTGGGTGAGTCTCTTAAGCAGCCTGCGATACTGGGTAGCTTCAATGTGAGCGGGCTT	883
Db	1284	CTTGGAATCATTCCTCAGCAGTGTAGATATTGGATGGCTTTTAAATATGAAGGGCCT	1343
Qy	884	GTACAGACTACTTCCACTGTGTATCTTCGATCAGCGGTCTCTAGAACTCTGCCAGGCC	943
Db	1344	GTCACTCTTATTTTGGGTGTGTTTCAGTGCATTAGTGGACCTCTCGGGAATGTACAGAACT	1403
Qy	944	CCAGGGAGCACGGCATGATG	963
Db	1404	CTTTGTCATGAGTTTGTG	1423

```

RESULT 13
US-09-949-016-132199
; Sequence 132199, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMERHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,769
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132199
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-132199

Query Match      12.7%; Score 137.4; DB 4; Length 601;
Best Local Similarity 58.5%; Pred: No. 2e-18;
Matches 300; Conservative 0; Mismatches 201; Indels 12; Gaps 3;

QY          61 CGGAGGTGCTGACCATCGCCCTCGCTGCATCCTGGGCCCTCATGCACCTGGGCCTTAC 120

```

Db	50	CGTGTGGTGGGCACAGCAGCTGTTTGGCCCTGGCAGTGTGGTGGCATCTCTGGGAGCCTAT	109
Qy	121	GCCTGCGGGGGTGCCTGCGCTCCGATCGCTACGGCCCTCTCTGGCCCTCTGAGCGG	180
Db	110	GTGACGGGCTACCAAGTTTCATCCACAGGAAAAGCATACCTGTCTCTTGGCCCTGTACGGC	169
Qy	181	GCCTTCTCTTTCACGCGCACCTGTGGTGGCGAGAGCCTCTTCCGCTACCTGGAGCACCGGCGG	240
Db	170	GCCATCTCTGGGCTGCACCTTGCTCATTTACAGAGCTTTTTCCTTCTCTGGAGCACCGGCGC	229
Qy	241	GTGCGCGCGCGCGCGGGGGCGCTGATGACGCCACACCGCGC---GCAGTGTGGCGCTG	297
Db	230	ATGCGAGCTGCGCGCGCAGGCCCTGAAAGTCCCTCCCGCGCGGGGCTCGGTGGCACATG	289
Qy	298	ACATCTCGCCTACACAGAGGAGACCGGCGTACTCTGGCCAGTGCCTTGGCGCTCGGCCCGC	357
Db	290	TGCAATTGCCGCRTACCAAGGAGGACCTGACTACTTTCGCGAAGTGCCTGCGCTCGGCCCGC	349
Qy	358	GCCTGTGCTACCGCGCGCGCGCTGCGGCTCTCATGTGTGTGATGGCAACCGCGCC	417
Db	350	CGCATCTCTTCCCTGAC-----CTCAAGTGGTCAATGGTGTGATGGCAACCGCCAG	403
Qy	418	GAGGACCTCTACATGTGTTCGACATGTTCCGCGAGGTCTT---CGCTGACGAGGACCCCGCC	474
Db	404	GAGGACGCTACATGCTGGACATCTTCCACGAGTGTCTGGCGCGCACCGAGCAGGCGCGC	463
Qy	475	ACGTAGCTGTGGAGCGGCAACTACCAACGACCTCTGGGAACCCCGCGCGCGCGCGCGGTG	534
Db	464	TTCTTTTGTGTGGCGCAGCACTTCCATGAGGACAGGCGAGGTGAGAGCGGCGCGCGCTG	523
Qy	535	GGCGCGGAGCCTATCGGAGGTGGAGGCGGAG	567
Db	524	CAGGAGGCATGACCTGTGCGGATGTGGTG	556
RESULT 14			
US-09-902-540-8819			
; Sequence 8819, Application US/09902540			
; Patent No. 6833447			
; GENERAL INFORMATION:			
; APPLICANT: Goldman, Barry S.			
; APPLICANT: Hinkle, Gregory J.			
; APPLICANT: Slater, Steven C.			
; APPLICANT: Wiegand, Roger C.			
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof			
; FILE REFERENCE: 38-10(15849)B			
; CURRENT APPLICATION NUMBER: US/09/902,540			
; CURRENT FILING DATE: 2001-07-10			
; PRIOR APPLICATION NUMBER: 60/217,883			
; PRIOR FILING DATE: 2000-07-10			
; NUMBER OF SEQ ID NOS: 16825			
; SEQ ID NO 8819			
; LENGTH: 6645			
; TYPE: DNA			
; ORGANISM: Myxococcus xanthus			
US-09-902-540-8819			
Query Match 7.6%; Score 82.8; DB 4; Length 6645;			
Best Local Similarity 44.8%; Pred. No. 1.6e-07;			
Matches 357; Conservative 0; Mismatches 437; Indels 2; Gaps 1;			
Qy	3	GAGACAGCAGGACGCGCCCAAGCCACTCTGTCAGCGCCGCGCTGCTCGCGGCTGGCCCG	62
Db	3823	GAGACGCGCCACCGCCGAGGTGCTTACCGCCGCGCGGCGGACCTCTTCGACGACCCG	3882
Qy	63	GAGGTTGCTGACCATCGCC---TTTCGCGCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC	120
Db	3883	AAGGACGCCCATCAGCGCTGTTTCGCGCTGGCGCGCTGCTCCCGGATGCTCGTC	3942
Qy	121	GCCTGCGGGGTGCGCTGAGCTCCGATGCGGTACCGGCTCTCTGGCCCTTGGCCCTTACGGG	180
Db	3943	ATCGACCGGCGCGTGGCGGGGTGTCAGCGGCTTGGAGCGTCAATGGGACCTGCTGGACGTG	4002

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 08:01:42 ; Search time 762.346 Seconds
(without alignments)
8450.560 Million cell updates/sec

Title: US-10-672-399-5

Perfect score: 1083

Sequence: 1 atgagacagcagcgccgccc.....tccaagtcgtactccgtga 1083

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:**
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1083	100.0	1083	18	US-10-672-399-5
2	953.4	88.0	1071	18	US-10-672-399-3
3	933	86.1	1737	18	US-10-672-399-1
4	933	86.1	2116	13	US-10-042-523-1
5	928	85.7	1065	18	US-10-672-399-7
6	677.8	62.6	1752	10	US-09-902-939-1
7	422	39.0	662	18	US-10-363-345A-13975
8	422	39.0	662	18	US-10-363-345A-13976
9	393.8	36.4	662	18	US-10-363-345A-13973
10	393.8	36.4	662	18	US-10-363-345A-13974
11	339.2	31.3	490	10	US-09-918-995-27210

12	273	25.2	1662	17	US-10-295-027-369	Sequence 369, App
13	273	25.2	1662	17	US-10-188-832-136	Sequence 136, App
14	250	23.1	1767	17	US-10-309-560-9	Sequence 9, Appli
15	236.4	21.8	1665	10	US-09-902-939-3	Sequence 3, Appli
16	179.4	16.6	1659	10	US-09-902-939-2	Sequence 2, Appli
17	179.4	16.6	4194	15	US-10-262-526-3	Sequence 3, Appli
18	153.6	14.2	3003	10	US-09-918-624B-64	Sequence 64, Appli
19	153.6	14.2	3003	15	US-10-262-526-1	Sequence 1, Appli
20	153.6	14.2	3387	15	US-10-084-1817-184	Sequence 184, App
21	153.6	14.2	3387	16	US-10-252-157-84	Sequence 84, Appli
22	153.6	14.2	4018	10	US-09-814-353-19942	Sequence 13942, A
23	100.2	9.3	793	9	US-09-910-943-552	Sequence 552, App
24	72	6.6	15559	18	US-10-646-664-1	Sequence 1, Appli
25	71	6.6	2182	18	US-10-437-963-62269	Sequence 62269, A
26	70	6.5	43058	9	US-09-954-456-292	Sequence 292, App
27	70	6.5	43058	9	US-09-954-456-529	Sequence 529, App
28	70	6.5	43058	9	US-09-880-107-3950	Sequence 3950, Ap
29	69.4	6.4	5784	14	US-10-152-886-64	Sequence 64, Appli
30	68.6	6.3	1755	18	US-10-437-963-78301	Sequence 78301, A
31	68.4	6.3	2214	15	US-10-156-761-2510	Sequence 2510, Ap
32	68.4	6.3	9025608	15	US-10-156-761-1	Sequence 1, Appli
33	67.6	6.2	1380	18	US-10-425-115-170699	Sequence 170699,
34	66.6	6.1	3957	16	US-10-300-562-193	Sequence 193, App
35	66.6	6.1	3957	16	US-10-237-551-193	Sequence 193, App
36	66.6	6.1	154746	10	US-09-827-688-8	Sequence 8, Appli
37	66.6	6.1	154746	10	US-09-827-688-8	Sequence 8, Appli
38	66.2	6.1	1614	9	US-09-976-740-45	Sequence 45, Appli
39	66.2	6.1	1614	13	US-10-023-529-45	Sequence 45, Appli
40	66.2	6.1	1614	13	US-10-023-523-45	Sequence 45, Appli
41	66.2	6.1	1614	17	US-10-616-187-45	Sequence 45, Appli
42	66.2	6.1	1614	17	US-10-671-242-45	Sequence 45, Appli
43	66.2	6.1	12425	9	US-09-976-740-50	Sequence 50, Appli
44	66.2	6.1	12425	13	US-10-023-529-50	Sequence 50, Appli
45	66.2	6.1	12425	13	US-10-023-523-50	Sequence 50, Appli

ALIGNMENTS

RESULT 1
US-10-672-399-5
; Sequence 5, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-5

Query Match 100.0%; Score 1083; DB 18; Length 1083;
Best Local Similarity 100.0%; Pred. No. 8.6e-257;
Matches 1083; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCGAGCCCGCGCTGCTCGGCGCTGCCC	60
Db	1	ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCGAGCCCGCGCTGCTCGGCGCTGCCC	60
Qy	61	CGGAGGGTGTGACCATCGCCTTGCCTGCTCATCTCGGCGCTCATGACCTGGGCGCTTAC	120
Db	61	CGGAGGGTGTGACCATCGCCTTGCCTGCTCATCTCGGCGCTCATGACCTGGGCGCTTAC	120
Qy	121	GCCGCCCGGGTGCCTGCGCTTCCGATCGCTACGCGCTTCTGGGCTTCTCGGCTCTTACGGG	180

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/10/042,523

APPLICATION NUMBER: US/10/042,523
FILING DATE: 19-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/635,552A
FILING DATE: 22-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Brook, David E.

REGISTRATION NUMBER: 22,592

REFERENCE/DOCKET NUMBER: LK95-07

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2116 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 36..1769

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-042-523-1

Query Match 86.1%; Score 933; DB 13; Length 2116;
Best Local Similarity 90.6%; Pred. No. 6.1e-220;
Matches 1048; Conservative 0; Mismatches 35; Indels 74; Gaps 2;

```
QY 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCGCGCGCTGCTCGGGCTGCGCC 60
DB 36 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCGCGCGCTGCTCGGGCTGCGCC 95
QY 61 CGGAGGCTGACCATCGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
DB 96 CGGAGGCTGACCATCGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 155
QY 121 GCGCGCGGGTGCGCTGCGCTCCGATCGCTACGCGCTTCTGCGCTTCTGCGCTTCTACGGG 180
DB 156 GCGCGCGGGTGCGCTGCGCTCCGATCGCTACGCGCTTCTGCGCTTCTGCGCTTCTACGGG 215
QY 181 GCCTTCTTTAGCGCACTGCTGCGCAGAGCCTTCTGCGCTACCTGAGAGCAGCGCGG 240
DB 216 GCCTTCTTTAGCGCACCCTGCTGCGCAGAGCCTTCTGCGCTACCTGAGAGCAGCGCGG 275
QY 241 GTGGCGCGCGCGCGCGCGCGCTGATGAGCCACCGCGCGCAGTGCTGCGCTGAC 300
DB 276 GTGGCGCGCGCGCGCGCGCGCTGATGAGCCACCGCGCGCAGTGCTGCGCTGAC 335
QY 301 ATCTCCGCTTACAGAGAGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 336 ATCTCCGCTTACAGAGAGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 395
QY 361 CTGCTGTATCCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
DB 396 CTGCTGTATCCCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 455
QY 421 GACCTCTACATGCTGATGCTTCCGCGAGGCTTCTGCTGAGCAGGACCCCGCCACCTAC 480
DB 456 GACCTCTACATGCTGATGCTTCCGCGAGGCTTCTGCTGAGCAGGACCCCGCCACCTAC 515
QY 481 GTGTGGGACGGCAATACCAACAGCCCTGGGAAACCCGCGCGCGCGCGCGCTGCGCGCC 540
DB 516 GTGTGGGACGGCAATACCAACAGCCCTGGGAAACCCGCGCGCGCGCGCGCTGCGCGCC 575
QY 541 GGAGCCTATCGGAGGCTGAGAGCGGAGGATCTCTGCGCGCGCTGCGAGGCGCTGCTG 600
DB 576 GGAGCCTATCGGAGGCTGAGAGCGGAGGATCTCTGCGCGCGCTGCGAGGCGCTGCTG 635
```

```
QY 601 AGGACTCGCAGTGGCTGCTGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCAATGTAC 660
DB 636 AGGACTCGCAGTGGCTGCTGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCAATGTAC 695
QY 661 ACAGCCTTCAAGCGCTCGGAGATTTCGGTGAATACGTCAGAGTCTGTGACTCGGACACA 720
DB 696 ACAGCCTTCAAGCGCTCGGAGATTTCGGTGAATACGTCAGAGTCTGTGACTCGGACACA 755
QY 721 AGGTTGAGACCCCATGGCACTGCTGAGCTGCTGCGGCTACTGGAGAGGACCCCGGGTA 780
DB 756 AGGTTGAGACCCCATGGCACTGCTGAGCTGCTGCGGCTACTGGAGAGGACCCCGGGTA 815
QY 781 GGGGCTGTTGGTGGGAGCTGCGGATCCTTAACCTCTGGACTCTGGGTCAGCTTCTCTA 840
DB 816 GGGGCTGTTGGTGGGAGCTGCGGATCCTTAACCTCTGGACTCTGGGTCAGCTTCTCTA 875
QY 841 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGAGCGGGCTTGTGAGACTACTTCCAC 900
DB 876 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGAGCGGGCTTGTGAGACTACTTCCAC 935
QY 901 TGTGTATCTGTCATCAGCGGCTCTAG----- 928
DB 936 TGTGTATCTGTCATCAGCGGCTCTAG----- 995
QY 929 -----AATCTCTGCCAGGCCCCAGGGAGCAGCGCATGATGCC 965
DB 996 CTTGAGGCTGGTACAAACAGAGTTCTGGGTACCACTGCTACTTTTGGGGATGACCGG 1055
QY 966 CTCATTCTCGCCCC-----CGTCAGGTACACTCCAGGTCCCGC 1006
DB 1056 CACCTCAACCAACCGCATGCTCAGCATGGGTATGCTACCAAGTACACCTCCAGGTCCCGC 1115
QY 1007 TGCTACTCAGAGAGCGCTTCTGCTGCGGTGGCTGAGCCAGCAGCAGCAGCTGCTCC 1066
DB 1116 TGCTACTCAGAGAGCGCTTCTGCTGCGGTGGCTGAGCCAGCAGCAGCAGCTGCTCC 1175
QY 1067 AAGTCGTACTTCCGTGA 1083
DB 1176 AAGTCGTACTTCCGTGA 1192
```

RESULT 5

US-10-672-399-7

; Sequence 7, Application US/10672399

; Publication No. US20050003368A1

; GENERAL INFORMATION:

; APPLICANT: University of Alberta

; TITLE OF INVENTION: Cancer Monitoring and Therapeutics

; FILE REFERENCE: A894635US

; CURRENT APPLICATION NUMBER: US/10/672,399

; PRIOR FILING DATE: 2003-09-25

; PRIOR APPLICATION NUMBER: US 60/472,401

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 7

; LENGTH: 1065

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-672-399-7

Query Match

Best Local Similarity 85.7%; Score 928; DB 18; Length 1065;

Matches 928; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCGCGCGCTGCTCGGGCTGCGCC 60
```

```
DB 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCGCGCGCTGCTCGGGCTGCGCC 60
```

```
QY 61 CGGAGGCTGCTGACCATCGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
```

```
DB 61 CGGAGGCTGCTGACCATCGCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
```

```
QY 121 GCGCGCGGGTGCCTGCGCTCCGATCGCTACGGCTCTCTGGCTTGGCTCTACGGG 180
Db 121 GCGCGCGGGTGCCTGCGCTCCGATCGCTACGGCTCTCTGGCTTGGCTCTACGGG 180
QY 181 GCCTTCTTTTACGCGACCTGTGGCGCAGAGCTCTTCGCTACTCTGAGCACCGGGCG 240
Db 181 GCCTTCTTTTACGCGACCTGTGGCGCAGAGCTCTTCGCTACTCTGAGCACCGGGCG 240
QY 241 GTGGCGCGGGCGCGCGGGCGCTGGATGAGCCACCGCGCGAGTGTGGCGCTGACC 300
Db 241 GTGGCGCGGGCGCGCGGGCGCTGGATGAGCCACCGCGCGAGTGTGGCGCTGACC 300
QY 301 ATCTCGGCTACACAGAGAACCCCGGCTACCTGGCGCAGTGCCTGGCGCTCGCGCGCC 360
Db 301 ATCTCGGCTACACAGAGAACCCCGGCTACCTGGCGCAGTGCCTGGCGCTCGCGCGCC 360
QY 361 CTGCTGTATACCCGCGCGGGTGGCGCTCTCATGTGGTGGTAAACCGCGCGGAG 420
Db 361 CTGCTGTATACCCGCGCGGGTGGCGCTCTCATGTGGTGGTAAACCGCGCGGAG 420
QY 421 GACCTCTACATGGTTCGACATGTTCCGCGAGGTCTTCGCTGACGAGAACCCCGCCACGTAC 480
Db 421 GACCTCTACATGGTTCGCGAGGTCTTCGCTGACGAGAACCCCGCCACGTAC 480
QY 481 GTGGGAGCGCACTACACAGCCCTGGGAAACCGCGCGCGCGCGCGCTGGCGCC 540
Db 481 GTGGGAGCGCACTACACAGCCCTGGGAAACCGCGCGCGCGCGCGCTGGCGCC 540
QY 541 GGAGCTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGCGCTGGTG 600
Db 541 GGAGCTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGCGCTGGTG 600
QY 601 AGGACTCGCAGGTGGCTGTGGCGCAGCGCTGGGGCGGCAAGCGCAGGTCTATGTAC 660
Db 601 AGGACTCGCAGGTGGCTGTGGCGCAGCGCTGGGGCGGCAAGCGCAGGTCTATGTAC 660
QY 661 ACAGCTTCAAGCGCTGGAGATTCGGTGGACTACGTGCAAGTCTGTGA CTCGACACA 720
Db 661 ACAGCTTCAAGCGCTGGAGATTCGGTGGACTACGTGCAAGTCTGTGA CTCGACACA 720
QY 721 AGGTTGGACCCCATGCTGCTGAGCTCGTGGCGGTACTGGAGGAGACCCCGGGTA 780
Db 721 AGGTTGGACCCCATGCTGCTGAGCTCGTGGCGGTACTGGAGGAGACCCCGGGTA 780
QY 781 GGGGCTGTGTGGGGAATGCGGATCTTAACCTCTGGAATCTTGGCTCAGCTTCTTA 840
Db 781 GGGGCTGTGTGGGGAATGCGGATCTTAACCTCTGGAATCTTGGCTCAGCTTCTTA 840
QY 841 AGCAGCTGCGATACCTGGTAGCCTTCAATGTGGAGCGGGCTTGTGACAGCTACTTCCAC 900
Db 841 AGCAGCTGCGATACCTGGTAGCCTTCAATGTGGAGCGGGCTTGTGACAGCTACTTCCAC 900
QY 901 TGTGTATCTGCATCAGCGGTCTCTAG 928
Db 901 TGTGTATCTGCATCAGCGGTCTCTAG 928
```

RESULT 6

US-09-902-939-1

```
; Sequence 1, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHazy
; APPLICANT: William Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/04020-USO
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1752
; TYPE: DNA
```

RESULT 7

```
; ORGANISM: mouse
US-09-902-939-1
Query Match 62.6%; Score 677.8; DB 10; Length 1752;
Best Local Similarity 84.5%; Pred. No. 3.4e-157;
Matches 792; Conservative 0; Mismatches 127; Indels 18; Gaps 2;
QY 10 CAGGACGCGCCCAAGCCCACTCTCTGACGCGCCCGCTGCTCGGCTTGGCCCGGAGGTTG 69
Db 7 CAGGACATGCAAAAGCCCTCAGAGCAGCGCTTCTGCTCTGGCTTGGCCAGGCGAGCA 66
QY 70 CTGACCATCGCTTCCGCTCTGCTCATCTTGGGCTCATGACCTGGGCTTACCGCGCGGG 129
Db 67 CTCACGATCATCTTTGGCTCTGCTCATCTGGGCTCATGACCTGGGCTTACCGCGAGGC 126
QY 130 GTGCGCTGGGCTCGGATCGCTACGCGCTCTCTGGGCTTCTGGGCTCTACGGGCTTCTT 189
Db 127 GTTCTCTGGCTTCAATGCTCTGATCTCTGGGCTTCTGGGCTCTATGGGCTATCTCTC 186
QY 190 TCAGCGCACCTGGTGGCGCAGAGCTCTTTCGCTTACCTGGAGCACCGCGGGTGGCGG 249
Db 187 AGCGCACACCTAGTGGCACAGAGCTCTTTCGCTTACCTGGAGCACCGAAGGCTGGCAG 246
QY 250 GGGGCGCG-----GGGGCGCTGGAATGCAAGCACCGCGGGCGAGTGGGG 294
Db 247 GCTGCGCGCGCTCTCTTGGCGAAGGGGCTCTTGGATGGCGGCTCTGCAAGCGCTGGCA 306
QY 295 CTGACCATCTCGGCTTACCGAGGAGCACCGCGCTACCTGGCGCAGTGTGGCGTCCGCC 354
Db 307 CTCACCATCTCAGCTTACCAAGAGATCTCCGCTTACCTGGCGCAGTGTGACCTCCGCC 366
QY 355 CGCGCTCTGCTGTATCCCGCGCGCGCTGCGCTCTCATGCTGGTGGATGGCAACCGC 414
Db 367 CGCGCTTGTCTATCCCGCACACGAGTTACGCTGCTCATGCTGGTGGACGCAACCGC 426
QY 415 GCGGAGGACCTCTATCATGCTGACATGTTTCGCGAGGCTTTCGCTGACGAGAACCCCGC 474
Db 427 GCTGAGGATCTGTATCATGCTGGACATGTTTCGAGAGTCTTTCGCGATGAGGACCCCGC 486
QY 475 AGTACGCTGTGGGACGGCAATACCAACGAGCCCTGGGAAACCGCGCGCG- --CGGCGG 531
Db 487 ACTATGCTGGGATGGCACTTACCATCAGCCCTGGGAAACCGAGAGCTACGGGCGCT 546
QY 532 GTGGCGCGCGGAGCTATCGGAGGTGGAGGAGGATCTCTGGCGGCTGGCAGTGGAG 591
Db 547 GTCGCTGAAGGTGCTTACCGGAGGTGGAGGCGGAGGACCCCGGCGGTTGGCGGTGGAG 606
QY 592 GCGCTGTGTGAGGACTCGCAGGTGCTGTGGCGCAGCGCTGGGGCGGGCAAGCGCGAG 651
Db 607 GCGCTGTGTGAGAACACGAGGTGCTGTGGCTCAGCGCTTGGGGCGGGCAACGCTGAG 666
QY 652 GTCATGTACACAGCTTCAAGCGCTCGAGATTCGGTGGGACTACGTCAGCTCTGTGAC 711
Db 667 GTCATGTACACAGCTTTCAGGGACTGGGCGACTCCGCTGGGACTAGTGGAGGCTGTGAC 726
QY 712 TCGGACACAAGGTTGGACCCCATGCTGCTGGAGCTCTGTCGGGCTACTTGGACGAGGAC 771
Db 727 TCAGACACAAGCTAGACCCCATGCTGCTGGAGCTTGTGGAGTGTGGATGAAGAC 786
QY 772 CCGCGGTAGGGCTGTGTGGGGAATGCGGAGTCTTAAACCTCTCTGGAATCTCTGGGTC 831
Db 787 CCGCGGTAGGGCTGTGTGGGGAATGCGGAGTCTTAAACCTCTCTGGAATCTCTGGGTC 846
QY 832 AGCTTCTTAAGCAGCTCGGATACCTGGGATGCTTCAATGTGGAGCGGGCTTGTGAGGC 891
Db 847 AGCTTCTTGAAGCAGCTTTCGATACCTGGGATGCTTCAATGTGGAGCGGGCTTGTGAGGC 906
QY 892 TACTTCCACTGTGTATCTGCTATCAGCGCTCTCTAG 928
Db 907 TACTTCCACTGTGTATCTGCTATCAGTGGTCTCTCTAG 943
```

US-10-363-345A-13975
 ; Sequence 13975, Application US/10363345A
 ; Publication No. US20040234960A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexander Olek
 ; APPLICANT: Kurt Berlin
 ; TITLE OF INVENTION: Method for determining the degree of methylation of defined
 ; FILE REFERENCE: E01/1227
 ; CURRENT APPLICATION NUMBER: US/10/363,345A
 ; CURRENT FILING DATE: 2003-03-03
 ; NUMBER OF SEQ ID NOS: 40712
 ; SEQ ID NO 13975
 ; LENGTH: 662
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; OTHER INFORMATION: CpG-island No: 13975
 US-10-363-345A-13975

Query Match 39.0%; Score 422; DB 18; Length 662;
 Best Local Similarity 77.3%; Pred. No. 2.8e-94;
 Matches 512; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

Qy	34	GCAGCCCGCGTGTCTCCGGCTGCGCCGGAGGGTGTGACATCGCCTTCGCCCTGCTC	93
Db	1	GTAGTTTGTGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTT	60
Qy	94	ATCCCTGGGCTCATGACCTGGGCTACCGCCGGGGTGGCGCTGGCTCGATCGCTAC	153
Db	61	ATTTTGGGTTTATGATTTGGGTTTACGTCGCGGGTGTGCTGGTTTCGATCGTTAC	120
Qy	154	GGCTCTCTGGGCTTCGGCTCTACGGGCTTCCTTTTCAGCGCACTCGTGGCGAGAGC	213
Db	121	GGTTTGTGTTTGGTTTACGGGTTTTCAGGTTTTCAGGTTTTCAGGTTTTCAGGTTT	180
Qy	214	CTCTTCGGTACCTGGAGACCGCGGGTGGCGGGTGGCGGGTGGCGGGTGGCGGGT	273
Db	181	TTTTTCGGCTATTTGGAGTATCGCGGGTGGCGGGTGGCGGGTGGCGGGTGGCGGGT	240
Qy	274	GCCACCGCGCAGTGTGCGCTGACCATCTCCGCTACAGGAGGACCCCGCGTACCTG	333
Db	241	GTTATCGCGTATTTGGGTTGTTGTTTTCGTTTATAGAGGATTCGCGTATTTG	300
Qy	334	CGCCAGTGCCTGGCTCCGCGCGCCCTGCTGTACCGCGCGCGCGCTGGCGTCTC	393
Db	301	CGTTAGTGTTCGCGTTCGTTTCGCTTTTGTGTTATTCGCGCGCGCTGGCGTTT	360
Qy	394	ATGTTGTGATGGCAACCGCGCGAGACCTCTACATGCTGTCATGTTTCGCGAGTCT	453
Db	361	ATGTTGTGATGGTGTATCGGCTCGAGGATTTTATATGCTGATGTTTCGCGAGTT	420
Qy	454	TTGCTGTGAGGACCCCGCAGTCTGTTGGAGGCTTACCTACCACTACCGCCCTGGAA	513
Db	421	TTGCTGTGAGGATTTCTGTTACGTACGTTGGAGGCTTATTTATTTAGTTTGGAA	480
Qy	514	CCCG	573
Db	481	TTGCG	540
Qy	574	GGCGCGCTGGAGTGTGAGGCTGTGAGGCTTCGACAGTGTGCTGGCGCGAGGCT	633
Db	541	GGCGCGTGTGAGTGTGAGGCTGTGAGGCTTCGACAGTGTGCTGGCGCGAGGCT	600
Qy	634	TGGCGCGCAGCGCGAGTGTACAGCTTCAAGCGCTCGGAGATTCGTTGGAC	693
Db	601	TGGCGCGTGTGAGGCTGTGAGGCTTCGACAGTGTGCTGGCGCGAGGCT	660
Qy	694	TA 695	
Db	661	TA 662	

RESULT 8
 US-10-363-345A-13976/c
 ; Sequence 13976, Application US/10363345A
 ; Publication No. US20040234960A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Alexander Olek
 ; APPLICANT: Kurt Berlin
 ; TITLE OF INVENTION: Method for determining the degree of methylation of defined
 ; FILE REFERENCE: E01/1227
 ; CURRENT APPLICATION NUMBER: US/10/363,345A
 ; CURRENT FILING DATE: 2003-03-03
 ; NUMBER OF SEQ ID NOS: 40712
 ; SEQ ID NO 13976
 ; LENGTH: 662
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; OTHER INFORMATION: CpG-island No: 13976
 US-10-363-345A-13976

Query Match 39.0%; Score 422; DB 18; Length 662;
 Best Local Similarity 77.3%; Pred. No. 2.8e-94;
 Matches 512; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

Qy	34	GCAGCCCGCGTGTCTCCGGCTGCGCCGGAGGGTGTGACATCGCCTTCGCCCTGCTC	93
Db	662	GTAGTTTGTGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTTGGTTT	603
Qy	94	ATCCCTGGGCTCATGACCTGGGCTACCGCCGGGGTGGCGCTGGCTCGATCGCTAC	153
Db	602	ATTTTGGGTTTATGATTTGGGTTTACGTCGCGGGTGTGCTGGTTTCGATCGTTAC	543
Qy	154	GGCTCTCTGGGCTTCGGCTCTACGGGCTTCCTTTTCAGCGCACTCGTGGCGAGAGC	213
Db	542	GGTTTGTGTTTGGTTTTCGGTTTTCAGCGGTTTTCAGCGTATTTTGGTGGCTAGGT	483
Qy	214	CTCTTCGGTACCTGGAGCACCGCGGGTGGCGGGTGGCGGGTGGCGGGTGGCGGGT	273
Db	482	TTTTTCGGTATTTGGAGTATCGCGGGTGGCGGGTGGCGGGTGGCGGGTGGCGGGT	423
Qy	274	GCCACCGCGCAGTGTGCGCTGACCATCTCCGCTACAGGAGGACCCCGCGTACCTG	333
Db	422	GTTATCGCGTATTTGGGTTGTTGTTTTCGTTTATAGAGGATTCGCGTATTTG	363
Qy	334	CGCCAGTGCCTGGCTCCGCGCGCCCTGCTGTACCGCGCGCGCGCTGGCGTCTC	393
Db	362	CGTTAGTGTTCGCGTTCGTTTCGCTTTTGTGTTATTCGCGCGCGCTGGCGTTT	303
Qy	394	ATGTTGTGATGGCAACCGCGCGAGACCTCTACATGCTGTCGATGTTTCGCGAGTCT	453
Db	302	ATGTTGTGATGGTATTCGCTCGAGGATTTTATATGCTGATATGTTTCGCGAGTT	243
Qy	454	TTGCTGTGAGGACCCCGCAGTCTGTTGGAGGCTTACCTACCACTACCGCCCTGGAA	513
Db	242	TTGCTGTGAGGATTTCTGTTACGTACGTTGGAGGCTTATTTATTTAGTTTGGAA	183
Qy	514	CCCG	573
Db	182	TTGCG	123
Qy	574	GGCGCGCTGGAGTGTGAGGCTGTGAGGCTTCGACAGTGTGCTGGCGCGAGGCT	633
Db	122	GGCGCGTGTGAGTGTGAGGCTGTGAGGATTCGTTAGGCTGCTGGCTGGCTAGGCT	63
Qy	634	TGGCGCGCAGCGCGAGTGTACAGCTTCAAGCGCTCGGAGATTCGTTGGAC	693
Db	62	TGGCGCGTGTGAGGCTGTGAGGCTTCGAGGCTTTCGAGGCTTCGTTGGAT	3

[illegible]

```

RESULT 11
US-09-918-995-27210
; Sequence 27210, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq. Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27210
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-27210

```

RESULT 12
US-10-295-027-369
; Sequence 369, Application US/10295027
; Publication No. US20030232350A1

```

: GENERAL INFORMATION:
: APPLICANT: Afar, Daniel
: APPLICANT: Aziz, Natasha
: APPLICANT: Ginsberg, Wendy M.
: APPLICANT: Gish, Kurt C.
: APPLICANT: Glynn, Richard
: APPLICANT: Hevezi, Peter A.
: APPLICANT: Mack, David H.
: APPLICANT: Murray, Richard
: APPLICANT: Watson, Susan R.
: APPLICANT: Eos Biotechnology, Inc.
: TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
: Methods of Screening for Modulators of Cancer
: FILE REFERENCE: 018501-012500US
: CURRENT APPLICATION NUMBER: US/10/295,027
: CURRENT FILING DATE: 2002-11-13
: PRIOR APPLICATION NUMBER: US 09/663,733
: PRIOR FILING DATE: 2000-09-15
: PRIOR APPLICATION NUMBER: US 60/350,666
: PRIOR FILING DATE: 2001-11-13
: PRIOR APPLICATION NUMBER: US 60/335,394
: PRIOR FILING DATE: 2001-11-15
: PRIOR APPLICATION NUMBER: US 60/332,464
: PRIOR FILING DATE: 2001-11-21
: PRIOR APPLICATION NUMBER: US 60/334,393
: PRIOR FILING DATE: 2001-11-29
: PRIOR APPLICATION NUMBER: US 60/340,376
: PRIOR FILING DATE: 2001-12-14
: PRIOR APPLICATION NUMBER: US 60/347,211
: PRIOR FILING DATE: 2002-01-08
: PRIOR APPLICATION NUMBER: US 60/347,349
: PRIOR FILING DATE: 2002-01-10
: PRIOR APPLICATION NUMBER: US 60/355,250
: PRIOR FILING DATE: 2002-02-08
: PRIOR APPLICATION NUMBER: US 60/356,714
: PRIOR FILING DATE: 2002-02-13
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 1386
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 369
: LENGTH: 1662
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-295-027-369

```

QY 418 GAGGACCTTACATGTCGACATGTTTCGCGAGGCTTTCGCTGACGAGAACCCCGCCACG 477
DB 382 GAGGACCCCTACATGTCGACATCTTCCACGAGTCTGCGCGGACACCGAGCAGGCGCGC 441
QY 478 TACGTGTGGACGGCAACTACACACAGCCCTGGGAACCCCGCGCGCGCGCGTGGGC 537
DB 442 TTCTTTGTGTGGCGCAGCAACTTCCA-----TGAGGCGAGCGGAGGTGAG 486
QY 538 GCCGAGCCCTATCGGAGGTGGAGCGGAGGATCTCTGGCGGTGGCACTGGAGGCGCTG 597
DB 487 ACGGAGCCAGCTCGAGGAGGCGATGACCGT-----GTGCGGATGTG 531
QY 598 GTGAGACTCTCGCAGGTGTGTGTGGCGCAGCGCTGGGCGGCGCAAGCGGAGGTGATG 657
DB 532 GTGCGGCGCAGCACCTTCTCGTGCATCATGCAAGTGGGAGGCAAGCGGAGGTGATG 591
QY 658 TACACAGCCCTCAAGCGCTCGGAGATTCCGAGATTCCGTTGAGTCTGACCTGCTCGGAC 717
DB 592 TACACGCGCTTCAAGCGCTTCCGCGGATTCGGTGGACTATCCAGGTGTGCGACTCTGAC 651
QY 718 ACAAGGTTGGACCCCATGCGACTGCTGGAGCTCGTGGCGGTACTGGACGAGCACCCCGG 777
DB 652 ACTGTGCTGATCCAGCTCGACCATCGAGATGCTTCGAGTCTTGGAGGAGATCCCCAA 711
QY 778 GTAGGGCTGTGTGGGAGCGTGGGATCTTAACTCTGAGCTCTGAGTCTGCGGTGACGTTTC 837
DB 712 GTAGGGGAGTGGGGAGATGTCCAGATCTTCAACAGTACGACTCATGGATTTCTTTC 771
QY 838 CTAAGCAGCTCGGATCTGGGTAGGCTTCAATGTGGAGCGGCGCTTGTGAGGCTACTTTC 897
DB 772 CTGACGAGCGTGGGTACTGGSATGGCTTCAACGTGGAGCGGCGCTGCCAGTCTACTTT 831
QY 898 CACTGTGTATCTGATCAGCGTCTCTAG 928
DB 832 GCCTGTGTGCACTGTATAGTGGGCCCTTGG 862

RESULT 13

US-10-188-832-136
; Sequence 136, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 136
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-188-832-136

Query Match

Best Local Similarity 25.2%; Score 273; DB 17; Length 1662;
Matches 532; Conservative 0; Mismatches 300; Indels 39; Gaps 4;

QY 61 CGGAGGGTGTGACCATCGCCCTTCCCTGTCTCATCTCTGGGCTCATGACCTGGGCGCTTAC 120
DB 28 CGTGTGTGGGCGACCCAGCGCTGTTTCCCTTGGCAGTGTGGTGGCATCTCTGGCAGCCTAT 87
QY 121 GCCGCGGGGTGCGCTTCCGATCGCTACGGGCTCTCTGGCTTCTGGGCTCTACGGG 180
DB 88 GTGACGGGTACCACTTATCCACACGGAAAGCACTACCTCTCTCTCTGGGCTGTACGGC 147
QY 181 GCTTCTCTTTTTCAGCGCACCTGTGTGGCGAGAGCCCTTCTCGCTATCTCTGAGACCCGCGG 240
DB 148 GGCATCTCTGGGCTGACCTGCTCATTCAGAGCCTTTTTCCTTCTGAGACCCGCGCGC 207
QY 241 GTGCGGCGGCGCGCGCGGCGCGCTGTGATGAGCAACCGCGC-----GCAGTGTGGGCTG 297
DB 208 ATGCGACGTGCGCGCGCAGGCGCTGAAGCTGCCCTCCCGCGCGGCGGCTCGGTGGCACTG 267
QY 298 ACCATCTCGGCTTACAGAGGAGACCCCGGTACTCTGCGCAGTGTGGGCTCTGGGCTCGCGCGC 357
DB 268 TGCATTGCGCGGTACAGGAGGACCTGACTACTTTCGCGCAAGTGTCTGCGCTCGGCGCCAG 327
QY 358 GCCCTGTGTATCCCGCGCGCGCGCTGCGCGTCTCTCATGTTGTGTGGATGGCAACCGCGCC 417
DB 338 CGCATCTCTCTCTCTGAC-----CTCAGGTGTGTGTTGTGATGGCAACCGCGCAG 381
QY 418 GAGGACCTTACATGTCGACATGTTCCGCGAGGTCTTCTGCTGACGAGACCCCGCGCAGC 477
DB 382 GAGGACGCTTACATGTCGACATCTTCCAGAGGTGTCTGGGCGGCAACCGAGCAGCGCGC 441
QY 478 TACGTGTGGAGCGGCAACTACACACAGCCCTGGGAAACCCCGCGCGCGCGGCTGGGCGC 537
DB 442 TTCTTTGTGTGGCGCAGCAACTTCCA-----TGAGGCGAGCGGAGGTGAG 486
QY 538 GCCGAGCGCTATCGGAGGTGGAGCGCGAGGATCTCTGGCGGCTGTGCACTGGAGGCGCTG 597
DB 487 ACGGAGCGCGCTCGAGAGGCGCATGACCGT-----GTGCGGATGTG 531
QY 598 GTGAGACTCGCAGGTGTGTGTGGCGCAGCGCTGGGCGGCGCAAGCGGAGGTGATG 657
DB 532 GTGCGGCGCAGCACCTTCTCGTGCATCATGCAAGTGGGAGGCAAGCGGAGGTGATG 591
QY 658 TACACAGCTTCAAGCGCTCGGAGATTGGTGGACTAGTGCAGGTCTGTGACTCGGAC 717
DB 592 TACACGCGCTTCAAGCGCTTCCGCGATTCGGTGGACTACATCCAGGTGTGCGACTCTGAC 651
QY 718 ACAAGGTTGGACCCCATGCGACTGCTGTGAGCTGTGCGGCTACTGGAACGAGCACCCCGG 777
DB 652 ACTGTGTGGATCCAGCGCTGACCATCGAGATGCTTCGAGTCTCTGAGGAGGATCCCCAA 711
QY 778 GTAGGGCTGTGTGGGAGCGTGGGATCTCTTAACTCTGAGCTCTGGGCTGAGCTTTC 837
DB 712 GTAGGGGAGTGGGGAGATGTCCAGATCTTCAACAGTACGACTCATGGATTTCTTTC 771
QY 838 CTAAGCAGCTCGGATCTGGGTAGGCTTCAATGTGGAGCGGCGCTTGTGAGGCTACTTTC 897
DB 772 CTGACGAGCGTGGGTACTGGSATGGCTTCAACGTGGAGCGGCGCTGCCAGTCTACTTT 831
QY 898 CACTGTGTATCTGATCAGCGTCTCTAG 928
DB 832 GCCTGTGTGCACTGTATAGTGGGCCCTTGG 862

RESULT 14

US-10-309-560-9
; Sequence 9, Application US/10309560
; Publication No. US20030235893A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: KUMARI, KSHAMA
; TITLE OF INVENTION: HYALURONAN SYNTHASES AND METHODS OF MAKING AND USING SAME
; FILE REFERENCE: 35541.082
; CURRENT APPLICATION NUMBER: US/10/309,560
; CURRENT FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 60/336,105

;
; PRIOR FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; TYPE: DNA
; ORGANISM: Xenopus laevis
US-10-309-560-9

```
Query Match      23.1%; Score 250; DB 17; Length 1767;
Best Local Similarity 55.8%; Pred. No. 5e-52;
Matches 503; Conservative 0; Mismatches 390; Indels 9; Gaps 1;

QY 62 GAGGGTGTGACCATGCGCTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAG 121
DB 80 GGAGGAAATATTTATTATCTTTGGTGGTCTATTAGCTACCAATTACAGCAGCTATG 139
QY 122 CGCCGGGGTGGCGTGGGCTCCGATCGCTAGCGCTCTGGGCTTCGGCCTCTACGGGG 181
DB 140 TGGCAGAGTTCAGGTCTCAACATGAAGCAATCTCTCTCCCTTGGGCTTTATGTC 199
QY 182 CTTCTCTTTACGGCACTGTGGCGCAGAGCTCTTCGCGTACCTGGAGCACCGCGGG 241
DB 200 TTGCAATGCTTCTCCACCTGATGATGACAGAGCTCTTTGGCTTCTCTGGAGATACGAGG 259
QY 242 TGGCGGGCGCGCGGGGGCGCTGGATGACGCCACCGCGGCTAGTGTGGGCTGACCA 301
DB 260 TAAATAAGAGTGA-----GCTTCTTTCGAGCTTTTAAAGAAGACAGTGGCTCTGACCA 310
QY 302 TCTCCGCTTACCAGAGGACCGCGGTACTCTGCGCAGTGTGCGTCCGCGCGCGGCC 361
DB 311 TTGCTGGGTATCAGAGAACCTTGATGATCTGTTGAATCTTGCAGATG 370
QY 362 TGCTGTACCGCGCGCGGGTGCCTGCTCATGAGTGTGATGGCAACCGCGCGAGG 421
DB 371 TGAATACCCCAAGATAAATCAAGATCATTTTGGTCAATCGATGGGAACACAGAGGATG 430
QY 422 ACCCTACATGTGTCACATGTTCCGCGAGGCTTTCGCTGACGAGAACCCCGCAGTACG 481
DB 431 ATGCTTACATGATGAGATGTTTCAAGACGTGTTCACCGTGAAGATGTAGGCACCTACG 490
QY 482 TGTGGGCGGCAACTACCAACAGCCCTGGGAACCGCGCGCGCGCGCGTGGCGCGC 541
DB 491 TATGGAAGGGAATTTACACACTGTTTAAAGCCTGAGGAGACCAATAGGGATCTGTG 550
QY 542 GAGCCTATCGGAGGTGAGGCGGAGGATCTTGGCGGGCTGCGCATGAGGCGCTGTGTA 601
DB 551 CTGAGGTTTCTAAGCCCTTGAATGAAGATGAAGGTATCAATATGTTGGAAGAACTTGT 610
QY 602 GGACTCGCAGGTGCGTGTGCGGCGCAGCGCTGGGCGGCGCAAGCGCAGTCTATGTACA 661
DB 611 GAAACAAGAGATGTGTGTGTCATATGCAACAGTGGGCGGGAAGAGAGTCTATGTACA 670
QY 662 CAGCCTTCAAGCGCTCGGAGATTGCGTGGACTAGTGCAGGTCTGTGACTCGGACACAA 721
DB 671 CAGCATTCAGGCCATTGGGACTTCTGTGGACTATGTACAGTCTGTGACTCGGACACCA 730
QY 722 GGTGGACCCCATGCTGCTGAGTCTGTGCGGTTACTGGAAGAGAACCCCGGGTAG 781
DB 731 AACTGGATGAATGCGCAACAGTGGAAATGGTGAAGTTCTGGAATCCAAATGACATGTACG 790
QY 782 GGGCTGTGGTGGGACGTGGGATCTTAAACCTCTGACTCTCTGGGTGAGTCTCTTAA 841
DB 791 CGCAGTGGGAGAGACGTTCCGATTTCTGAACCTTATGATTCCTTATATGTTTCTATGA 850
QY 842 GCAGCTCGGATCTGCGGTAGCTTCAATGTGGAGCGGGCTTGTTCAGAGTACTTCCACT 901
DB 851 GCAGCTCGGTACTGGATGCGGTTTAACTGTGGAGGGGCTTGCAGTCTTACTTCGACT 910
QY 902 GTGTATCTGTCATGACGCGGTCTCTTAGAATCTGTGCGGACGCGCCAGGGAGCACGCGATGA 961
DB 911 CGGTGTCCTGTATAAGTGGACCTCTGTGGATGTACCGGAACAAATCTCCAGGTTGTTT 970
```

QY 962 TG 963
DB 971 TG 972

RESULT 15

US-09-902-939-3
; Sequence 3, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHazy
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/0020-US0
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 1665
; TYPE: DNA
; ORGANISM: mouse
US-09-902-939-3

```
Query Match      21.8%; Score 236.4; DB 10; Length 1665;
Best Local Similarity 58.5%; Pred. No. 1.1e-48;
Matches 511; Conservative 0; Mismatches 321; Indels 42; Gaps 4;

QY 61 CGGAGGGTGTGTGACCATGCGCTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
DB 28 CGTGTGTGGGCAACCACTGTGTTGGCTGTGTGTGTGTGGAGGATCTCTGGCGGCTAT 87
QY 121 GCCCGCGGGGTGCGCTGGGCTTCGATCGCTACCGGCTCTCTGGCTTCGGCTCTTACGGG 180
DB 88 GTGACAGGCTACCAAGTTTATCCACACAGAAAGACACTACTGTCTTTGGGCTCTACGGT 147
QY 181 GCCTTCCTTTTCAGCGCACCTGTGGCGCAGAGCTCTTCGGTACCTGGACACCGGGCGG 240
DB 148 GCCATCTTGGGTCTACATCTGCTCATCAGAGCTGTTTGCCTTCTTGGAGCACCGTGA 207
QY 241 GTGCGCGCGCGCGCGCG-----GGGCGCGCTGTGATGAGCCACCGCGCGCAGTGTGGCG 294
DB 208 ATGCGCAGGCGAGGCGCGCCCTCAAGTGCATGCTCCACAGAGGTCCGCTTCAGTGGCA 267
QY 295 CTGACCATCTCCGCTTACAGAGGACCCGCGTACCTGCGCCAGTGTGCTGGCGCTCCGCC 354
DB 268 CTCTGCAATGCTGCTTACCAAGAGGACCCCGAATACCTGCGCAAGTGTCTTCGCTCAGCT 327
QY 355 CGCGCCTGTGTATACCGCGCGCGCTGCGCGCTCTCATGTGTGGATGGCAACCGC 414
DB 328 CAGCGCATGCTTTCCTTCCAAAC-----CTCAAGGTGTGTGATGGTCAATGCG 381
QY 415 GCCGAGGACCTCTACATGTGTGACATGTTCCGAGGTCTTCGCTGACGAGACCCCGCC 474
DB 382 CAGGAAGATACCTACATGTTGGACATCTTCCATGAGGTGTGGTGGCACTGAGCAGACT 441
QY 475 ACGTACTGTGGAGCGGCAACTACACAGCCCTTGGGAAACCCGCGCGCGCGCGCGGTG 534
DB 442 GGCCTTCTTGTGTGGCGTAGCAATTTCCAT-----GAGGCGGGTGAAGGA 486
QY 535 GCGCGCGGACCTATCGGAGGTGGAGCGGAGGATCTCTGGCGGCTGGCAGTGGAGCGG 594
DB 487 GAGACAGAGGCGACCTGTCAGGAAGGATGGAGCGT-----GTCCGAGCT 531
QY 595 CTGTGTAGGACTCTCAGGTGTGTGTGTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTC 654
DB 532 GTGTGTGGGCGCAGCACTTCTCATGCAATCATCAGAAAGTGGGGGGCAAGCGTGAAGTC 591
QY 655 ATGTACACAGCCTTTAAAGCGCTCGGAGATTCGGTGAAGTACGTGCAAGGTCTGTGACTCG 714
DB 592 ATGTACACAGCCTTTAAAGCGCTTGGCAACTCAGTGGACTACATCCAGGTTGTGACTCT 651
QY 715 GACACAGGTTGGACCCCATGGCACTGTCTGGAGTCTCTGGGGTACTTGGACGAGGACCCC 774
```


THIS PAGE BLANK (USTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 13:46:32 ; Search time 101.25 Seconds
(without alignments)
1375.148 Million cell updates/sec

Title: US-10-672-399-6
Perfect score: 1864
Sequence: 1 MRQDAPKPTPAARCSGLA.....ALVLPAVAPADTLVQVLP 360

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1605	86.1	578	2 AAW26765	AAW26765 Human hya
2	1512	81.1	582	3 AAY78135	AAY78135 Recombina
3	1512	81.1	583	2 AAW30704	AAW30704 Mouse hya
4	1512	81.1	583	2 AAY32503	AAY32503 Hyaluro
5	1512	81.1	583	3 AAY68491	AAY68491 Mouse hya
6	1512	81.1	583	3 AAB09948	AAB09948 Murine HA
7	1512	81.1	583	6 AAB96028	ABP96028 Mouse hya
8	1512	81.1	583	7 AAE39152	AAE39152 Mouse hya
9	1425	76.4	543	2 AAW36503	AAW36503 Human hya
10	1419.5	76.2	584	3 AAY78132	AAY78132 Recombina
11	1419	76.1	582	3 AAY78138	AAY78138 Recombina
12	1419	76.1	583	3 AAY78131	AAY78131 Recombina
13	1238	66.4	563	3 AAY78129	AAY78129 Recombina
14	1000.5	53.7	577	3 AAY78128	AAY78128 Recombina
15	992	53.2	573	3 AAY78127	AAY78127 Recombina
16	876.5	47.0	552	3 AAY78133	AAY78133 Recombina
17	876.5	47.0	553	3 AAY78136	AAY78136 Recombina
18	811	43.5	552	7 ADC49213	ADC49213 Rabbit hy
19	804.5	43.2	588	7 ADD93928	ADD93928 Xenopus I
20	800	42.9	554	3 AAW50010	AAW50010 Murine hy
21	800	42.9	554	3 AAY68493	AAY68493 Mouse hya
22	800	42.9	554	6 AAB96030	ABP96030 Mouse hya
23	800	42.9	554	7 AAE39154	AAE39154 Mouse hya
24	800	42.9	554	8 ADJ76214	ADJ76214 Marker Ge
25	790.5	42.4	552	3 AAY78140	AAY78140 Recombina

26	790.5	42.4	553	6 ABR48209	ABR48209 Human bla
27	790.5	42.4	553	6 ABUS6490	ABUS6490 Lung canc
28	790.5	42.4	553	6 ABUS6686	ABUS6686 Lung canc
29	790.5	42.4	553	7 ADN39052	ADN39052 Cancer/an
30	790.5	42.4	553	8 ADJ75425	ADJ75425 Marker ge
31	790.5	42.4	553	8 ADN05116	ADN05116 Antipso
32	787.5	42.2	552	3 AAY78139	AAY78139 Recombina
33	787.5	42.2	552	3 AAY78142	AAY78142 Recombina
34	784	42.1	553	3 AAY78134	AAY78134 Recombina
35	783.5	42.0	552	2 AAW50009	AAW50009 Murine hy
36	783.5	42.0	552	3 AAY78141	AAY78141 Recombina
37	783.5	42.0	552	3 AAY68492	AAY68492 Mouse hya
38	783.5	42.0	552	5 ABB57265	ABB57265 Mouse isc
39	783.5	42.0	552	6 AAB96029	ABP96029 Mouse hya
40	783.5	42.0	552	7 ADA49686	ADA49686 Mouse hya
41	783.5	42.0	552	7 ADA49684	ADA49684 Human hya
42	783.5	42.0	552	7 AAE39153	AAE39153 Mouse hya
43	783.5	42.0	552	7 ADC59297	ADC59297 Human HAS
44	783.5	42.0	552	7 ADC59299	ADC59299 Mouse HAS
45	783.5	42.0	552	7 ADD48818	ADD48818 Human Pro

ALIGNMENTS

RESULT 1
AAW26765
ID AAW26765 standard; protein; 578 AA.
XX AC AAW26765;
XX DT 21-MAY-1998 (first entry)
XX DE Human hyaluronan synthase.
XX KW Hyaluronan synthase; HAS gene; human; hyaluronic acid; cell adhesion;
XX KW wound healing; vulnery; tissue repair; scar; keloid; therapy.
XX OS Homo sapiens.
XX FH Key
FT Binding-site Location/Qualifiers
FT Modified-site /note= "putative hyaluronan binding site"
FT Modified-site /note= "consensus phosphorylation sequence for protein kinase C"
FT Modified-site /note= "consensus phosphorylation sequence for protein kinase C"
FT Modified-site /note= "consensus phosphorylation sequence for protein kinase C"
FT Modified-site /note= "consensus phosphorylation sequence for cAMP-dependent protein kinase"
XX WO9740174-A1.
XX PD 30-OCT-1997.
XX PF 17-APR-1997; 97WO-US006350.
XX PR 22-APR-1996; 96US-00635552.
XX PA (LEUK-) LEUKOSITE INC.
XX PI Brieskin MJ;
XX DR WPI; 1997-549359/50.
XX DR N-PSDB; AAT99541.
XX PT Human hyaluronan synthase - useful for recombinant production of
PT hyaluronic acid for wound healing, tissue repair and reducing

PT hypertrophic scar and keloid formation.

XX Claim 5; Page 38-40; 58pp; English.

CC This protein comprises human hyaluronan synthase (HAS), an enzyme
 CC involved in the synthesis of hyaluronan (hyaluronic acid) and which has
 CC the ability to confer cell adhesion by the lymphocyte receptor CD44. The
 CC amino acid sequence was deduced from clone 30C (see AAT99541), which was
 CC obtained from a human mesenteric lymph node expression library using an
 CC expression cloning system developed to isolate cDNA clones that encode
 CC proteins that confer adhesion of the murine T cell lymphoma TK1. The
 CC isolated clone can be utilised in a claimed method for producing HAS in
 CC host cells. Such host cells are used in a claimed method for the
 CC production of hyaluronan. Hyaluronan is useful for wound healing and
 CC tissue repair, and can reduce or prevent hypertrophic scars and keloid
 CC formation. It is also used in eye surgery as a replacement for vitreous
 CC fluid

XX Sequence 578 AA;

Query Match 86.1%; Score 1605; DB 2; Length 578;
 Best Local Similarity 96.0%; Pred. No. 1.8e-170;
 Matches 312; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

QY 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYG 60
 DB 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYG 60
 QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILASARA 120
 DB 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILASARA 120
 QY 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVDGNHYPWEPAAAGAVGA 180
 DB 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVDGNHYPWEPAAAGAVGA 180
 QY 181 GAYREVEAEDPGLAVEALVTRRCVCAQRWGKREVMYTAFAKALGDSVDYVQVCDSDT 240
 DB 181 GAYREVEAEDPGLAVEALVTRRCVCAQRWGKREVMYTAFAKALGDSVDYVQVCDSDT 240
 QY 241 RLDPMALLELVRVLDDEDPRVGAGVDVRLNPLDSWVSFLSSRLRYWVAFNVERACQSYFH 300
 DB 241 RLDPMALLELVRVLDDEDPRVGAGVDVRLNPLDSWVSFLSSRLRYWVAFNVERACQSYFH 300
 QY 301 CVSCISGPLESCPGRHAMMPSFL 325
 DB 301 CVSCISGPLESCPGRHAMMPSFL 325

RESULT 2

AAV78135

ID AAV78135 standard; protein; 582 AA.

XX AAV78135;

XX 27-APR-2000 (first entry)

DT Recombinant chimeric hyaluronate synthase modified protein #9.

DE Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;

KW research reagent; biochemical research; medical development; chimeric.

XX Mus sp.

OS Synthetic.

OS Chimeric.

XX JP2000004886-A.

PN 11-JAN-2000.

FD 24-JUN-1998; 98JP-00193788.

XX 24-JUN-1998; 98JP-00193788.

XX (SEGG) SEIKAGAKU KOGYO CO LTD.

XX WPI; 2000-140125/13.

XX A hyaluronate synthase modified protein - useful as a research reagent
 PT for biochemical research and medical development.

XX Claim 10; Page; 30pp; Japanese.

XX The present invention describes a recombinant protein which consists of
 CC three continuous regions (N-terminal region, internal region and C-
 CC terminal region) where one or two regions among the above three regions
 CC is selected from the three hyaluronate synthase (HAS) modified proteins
 CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
 CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
 CC is useful as a research reagent for biochemical research and medical
 CC development. The invention provides a HAS modified protein of increased
 CC or lowered activity. AAV78127 to AAV78142 represent specifically claimed
 CC examples of recombinant proteins from the present invention. N.B. The
 CC present sequence is not given in the present specification, but is
 CC derived from sequences given as specified in the claim

XX Sequence 582 AA;

Query Match 81.1%; Score 1512; DB 3; Length 582;
 Best Local Similarity 90.6%; Pred. No. 5.1e-160;
 Matches 298; Conservative 5; Mismatches 16; Indels 10; Gaps 3;

QY 3 QODAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYGAF 62
 DB 2 QODAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYGAF 61
 QY 63 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILAS 117
 DB 63 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILAS 121
 QY 118 ARALLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVDGNHYPWEPAAAGAVGA 176
 DB 122 ARALLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVDGNHYPWEPAAAGAVGA 181
 QY 177 AVGAGAYREVEAEDPGLAVEALVTRRCVCAQRWGKREVMYTAFAKALGDSVDYVQVVC 236
 DB 182 AVGAGAYREVEAEDPGLAVEALVTRRCVCAQRWGKREVMYTAFAKALGDSVDYVQVVC 241
 QY 237 DSDTRLDPMALLELVRVLDDEDPRVGAGVDVRLNPLDSWVSFLSSRLRYWVAFNVERACQ 296
 DB 242 DSDTRLDPMALLELVRVLDDEDPRVGAGVDVRLNPLDSWVSFLSSRLRYWVAFNVERACQ 301
 QY 297 SYFHCVSCISGPLESCPGRHAMMPSFL 325
 DB 302 SYFHCVSCISGPLESCPGRHAMMPSFL 326

RESULT 3

AAW30704

ID AAW30704 standard; protein; 583 AA.

XX AAW30704;

XX 17-OCT-2003 (revised)

DT 19-DEC-1997 (first entry)

XX Mouse hyaluronate synthase enzyme.

XX Hyaluronate synthase; hyaluronic acid; polysaccharide.

XX Mus sp; (strain FM3A).

XX JP09224674-A.

XX 02-SEP-1997.

Dd	242	DSDFRLDPMALLELVRLVLDDEPRVGAVGGDVIRILNPLPDSWVSFLSSLRYYWAFNVERACQ	301
Qy	297	SYFHCVCISGPLESCPGPREHAMMPSFL	325
Dd	302	SYFHCVCISGGLPL---GLYRNLLQQFL	326
RESULT 8			
AAE39152	ID	AAE39152 standard; protein; 583 AA.	
XX	AC	AAE39152;	
XX	DT	18-DEC-2003 (first entry)	
XX	DE	Mouse hyaluronan synthase (HAS) 1.	
XX	KW	Mouse; hyaluronic acid; HA; hyaluronan synthase; HAS; dry eye syndrome;	
XX	KW	eye cell; osteoarthritis; gene therapy; enzyme.	
XX	OS	Mus sp.	
XX	PN	US2003087850-A1.	
XX	PD	08-MAY-2003.	
XX	PF	10-JUL-2001; 2001US-00902939.	
XX	PR	10-JUL-2001; 2001US-00902939.	
XX	PA	(DEHA)/ DEHAZYA P.	
XX	PA	(CHEN/) CHEN W.	
XX	PI	Dehazya P., Chen W;	
XX	DR	WFI; 2003-755151/71.	
XX	DR	N-PSDB; AAD59442.	
PT	CC	Dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate	
PT	CC	comprising derivatized hyaluronic acid cross linked to nucleic acid	
PT	CC	encoding hyaluronan synthase useful for treating dry eye syndrome.	
XX	PS	Claim 18; Page 15-16; 31pp; English.	
XX	CC	The invention relates to dihydrazide derivatised hyaluronic acid (HA) /	
XX	CC	nucleic acid bioconjugate comprising derivatised HA cross linked to	
XX	CC	nucleic acid encoding hyaluronan synthase (HAS). The invention is useful	
XX	CC	for transfected an eye cell of an individual. It is useful for treating	
XX	CC	dry eye syndrome and osteoarthritis of the particular joints. The	
XX	CC	invention is also useful in gene therapy. The present sequence is mouse	
XX	CC	HAS1 enzyme	
SQ		Sequence 583 AA;	
Qy	Query Match	81.1%; Score 1512; DB 7; Length 583;	
Dd	Best Local Similarity	90.6%; Pred. No. 5.1e-160;	
Qy	Matches 298; Conservative	5; Mismatches 16; Indels 10; Gaps 3	
Qy	3 QDAPKTPAARCSGLARRVLITAFALLIIGLMTWAYAAGVPPLASDRYGLLAFLGYGAF	62	
Dd	2 RDMPKPESEARCCSGLARRALTIIFALLIIGLMTWAYAAGVPPLASDRYGLLAFLGYGAF	61	
Qy	63 LSAHLVAQSIPAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCCLAS	117	
Dd	62 LSAHLVAQSIPAYLEHRRVAAAARRSLAKGPLDAATARSVALTISAYQEDPAYLRQCCLTS	121	
Qy	118 ARALLYPRABRLVLMVVVDGNRAEDLYVMVDMREFVFADEDPATYVMDGNYNHPWEPA-AAAG	176	
Dd	122 ARALLYPHTRLRVLMVVVDGNRAEDLYXVDMREFVFADEDPATYVMDGNYNHPWEPAEATG	181	
Qy	177 AVGAGAYREVEADPPGRILAVALVRTRCVCVAORWGKRREWMYTAFKALGSDVDYVQVC	236	
Dd	182 AVGEGAYREVEADPPGRILAVALVRTRCVCVAORWGKRREWMYTAFKALGSDVDYVQVC	241	

	Mus sp.	
XX	WO200306068-A1.	
XX	23-JAN-2003.	
XX	10-JUL-2001; 2001WO-US021785.	
XX	10-JUL-2001; 2001WO-US021785.	
XX	(CLEA-) CLEAR SOLUTIONS BIOTECH INC.	
XX	Dahzya P, Chen W;	
XX	WPI; 2003--221664/21.	
DR	N-PSTDB; ABZ76734.	
PT	Noel dihydrazone derivatized hyaluronic acid/nucleic acid bioconjugate for treating dry eye syndrome, has derivatized hyaluronic acid crosslinked to nucleic acid encoding protein with hyaluronan synthase activity.	
PS	Claim 20; Page 58-59; 62pp; English.	
CC	The present invention describes a dihydrazide derivatised hyaluronic acid (HA)/nucleic acid (NA) bioconjugate (I), with derivatised HA linked to NA encoding protein with hyaluronan synthase (HAS) activity, where NA has sequence with 79-85% sequence identity (SI) to a sequence of 1752. 1659 or 1665 nucleotides (see ABZ76734 to ABZ76736) or a sequence encoding proteins with 66-90% SI to a sequence of 583, 552 or 554 amino acids (see ABP96028 to ABP96030). (I) has ophthalmological, osteopathic and antiarthritic activities, and can be used in gene therapy and as an inhibitor of angiogenesis, and as an inducer of expression of (HA) in human corneal epithelial cell. (I) is useful for transfecting a cell of an eye with nucleic acid, where the nucleic acid comprises (ABZ76734 to ABZ76736), by contacting the cell with (I) comprising the nucleic acid. (I) is useful for treating dry eye syndrome in an individual. (I) is useful in gene therapy applications for the treatment of a variety of medical conditions including dry eye syndrome or other medical conditions where an increase in the production of (HA) in the eye would be therapeutic (e.g., osteoarthritis of the articular joints). (I) is also useful for inhibiting angiogenesis for the treatment of macular degeneration or genes related to lipid biosynthesis that helps to restore the lipid component of the tear film, and as reagents for in vitro transformation of any cell, preferably a eukaryotic cell, more preferably a human eye cell. The present sequence represents mouse HAS1 which is used in the exemplification of the present invention	
SQ	Sequence 583 AA;	
	Query Match 81.1%; Score 1512; DB 6; Length 583; Best Local Similarity 90.6%; Pred. No.5.le-160; Matches 298; Conservative 5; Mismatches 16; Indels 10; Gaps 3;	
Qy	3 QDAPKTPPAARRCSGLARRVLTATFALLILGLMTWAYAAAGVPPLASRDYRGLLAGLYGAF 62 : :	
Db	2 RQDMPKSEARCCSGLARALTIFALLIIIGLMTWAYAAAGVPPLASRDYRGLLAGLYGAF 61 : :	
Qy	63 LSAHLVAQSUFAYLEHRVAAAAAR-----GPLDAATARSVALTISAVGEDPAYLRQCilas 117 : :	
Db	62 LSAHLVAQSUFAYLEHRVAAAARSRLAKGPLDAATARSVALTISAVGEDPAYLRQCILTS 121 : :	
Qy	118 ARALLYPRALRVLMVDGNRAEDLVMDMFEVFADDPATYYWDNGNYHPWEPA-NAG 176 : :	
Db	122 ARALLYPHTRLRVLVMVDGNRAEDLVMDMFEVFADDPATYYWDNGNYHQPWFPEATG 181 : :	
Qy	177 AVGAGAREVEAADPGRLAVEALVRTRCVCVAORWGGRKMVTAFKALGDSDYDVGVOC 236 : :	
Db	182 AVGEGAREVEAADPGRLAVEALVRTRCVCVAORWGGRKMVTAFKALGDSDYDVGVOC 241 : :	
Qy	237 DSDTRLDPALLEVRVLDEDDPVGA VG GD VRIINLPDSWSYFLSSIRRYWFAFNVERACQ 296 : :	

Db 122 ABALYPPHTRLRVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAATG 181
Qy 177 AVGAGAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMYT-APKALGDSVDYVQV 235
Db 182 AVGAGAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGNSVDYVQV 241
Qy 236 CSDTRLDPMALLEVRVLDDEPRVGA VGDVRIINPLDSWVSFLSSRYWYAFNVERAC 295
Db 242 CSDTVLDPACTIEMLRVLEEDPQVGGVGDVQIILNKYDSWISFLSSRYWYAFNVERAC 301
Qy 296 QSYFHCVSCISGPLESCPGPREHAMMPSFL 325
Db 302 QSYFGVCQCISGPL-----GMYNRLSLQOFL 327

RESULT 11
AA78138
ID AAY78138 standard; protein; 582 AA.
XX
AC AAY78138;
XX
DT 27-APR-2000 (first entry)
XX
DE Recombinant chimeric hyaluronate synthase modified protein #12.
XX
KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX
OS Mus sp.
OS Synthetic.
OS Chimeric.
XX
PN JP2000004886-A.
XX
PD 11-JAN-2000.
XX
PF 24-JUN-1998; 98JP-00193788.
XX
PR 24-JUN-1998; 98JP-00193788.
XX
PA (SEKG) SEIKAGAKU KOGYO CO LTD.
XX
DR WPI; 2000-140125/13.
XX
PT A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.
XX
PS Claim 13; Page; 30pp; Japanese.

CC The present invention describes a recombinant protein which consists of
CC three continuous regions (N-terminal region, internal region and C-
CC terminal region) where one or two regions among the above three regions
CC is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. AAY78127 to AAY78142 represent specifically claimed
CC examples of recombinant proteins from the present invention. N.B. The
CC present sequence is not given in the present specification, but is
CC derived from sequences given as specified in the claim
XX
SQ Sequence 582 AA;

Query Match 76.1%; Score 1419; DB 3; Length 582;
Best Local Similarity 83.9%; Pred. No. 1.4e-149;
Matches 276; Conservative 18; Mismatches 25; Indels 10; Gaps 3;

Qy 3 QODAKPTPAARRCCSGLARRVLTIAFALLILGLMTWAYAGVPLASDRYGLLAFLYGA 62
Db 2 RQDMKPSEARCCSGLARRALTITFALLILGLMTWAYAGVPLASDRYGLLAFLYGA 61

Qy 63 LSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCLAS 117
Db 62 LSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCLTS 121
Qy 118 ARALYPPARLRVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPA-AG 176
Db 122 ARALYPPHTRLRVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAATG 181
Qy 177 AVGAGAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMYTAFKALGDSVDYVQV 236
Db 182 AVGAGAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQV 241
Qy 237 DSDTRLDPMALLEVRVLDDEPRVGA VGDVRIINPLDSWVSFLSSRYWYAFNVERACQ 296
Db 242 DSDTMLDPASSVMVVKVLEEDPMVGGVGDVQIILNKYDSWISFLSSRYWYAFNVERACQ 301
Qy 297 SYFHCVSCISGPLESCPGPREHAMMPSFL 325
Db 302 SYFGVCQCISGPL-----GMYNRLSLHEFV 326

RESULT 12
AA78131
ID AAY78131 standard; protein; 583 AA.
XX
AC AAY78131;
XX
DT 27-APR-2000 (first entry)
XX
DE Recombinant chimeric hyaluronate synthase modified protein #5.
XX
KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX
OS Mus sp.
OS Synthetic.
OS Chimeric.
XX
PN JP2000004886-A.
XX
PD 11-JAN-2000.
XX
PF 24-JUN-1998; 98JP-00193788.
XX
PR 24-JUN-1998; 98JP-00193788.
XX
PA (SEKG) SEIKAGAKU KOGYO CO LTD.
XX
DR WPI; 2000-140125/13.
XX
PT A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.
XX
PS Claim 6; Page; 30pp; Japanese.

CC The present invention describes a recombinant protein which consists of
CC three continuous regions (N-terminal region, internal region and C-
CC terminal region) where one or two regions among the above three regions
CC is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. AAY78127 to AAY78142 represent specifically claimed
CC examples of recombinant proteins from the present invention. N.B. The
CC present sequence is not given in the present specification, but is
CC derived from sequences given as specified in the claim
XX
SQ Sequence 583 AA;

Query Match 76.1%; Score 1419; DB 3; Length 583;
Best Local Similarity 83.9%; Pred. No. 1.4e-149;
Matches 276; Conservative 18; Mismatches 25; Indels 10; Gaps 3;

```

QY 3 QODAKPTPAARCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFLGYGAF 62
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2 RODMPKPSAARCCSGLARRAIIIFALLILGLMTWYAAGVPLASDRYGLLAFLGYGAF 61
QY 63 LSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCILAS 117
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
62 LSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCILTS 121
QY 118 ARALYPPARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNYHQWPEPA-AAG 176
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
122 ARALYPPHTRLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNYHQWPEPAEATG 181
QY 177 AVGAGAYREVEAEDPGRLAVEALVTRRCVCVQAORWGKREVMYTAFKALGDSVDYVQVC 236
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 AVGEGAYREVEAEDPGRLAVEALVTRRCVCVQAORWGKREVMYTAFKALGRSVDYVQVC 241
QY 237 DSDTRLDPMALLELRVLDDEDPVGAAGVDVRLNPLDSWVSFLSSLRVWVAFNVERACQ 296
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
242 DSDTMDLPASSVEMVKVLEEDPMVGCGVDVQVILNKYDSWISFLSSRVWVAFNVERACQ 301
QY 297 SYFHCVCISGPLESCPGPREHAMMPSFL 325
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
302 SYFGVCVQCISGPL-----GWRNSLLHEFV 326

```

RESULT 13

```

AAV78129
ID AAV78129 standard; protein; 563 AA.
XX
AC AAV78129;
XX

```

```

DT 27-APR-2000 (first entry)
XX

```

```

DE Recombinant chimeric hyaluronate synthase modified protein #3.
XX

```

```

KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX

```

```

OS Mus sp.
XX

```

```

OS Synthetic.
XX

```

```

OS Chimeric.
XX

```

```

FN JP2000004886-A.
XX

```

```

PD 11-JAN-2000.
XX

```

```

PF 24-JUN-1998; 98JP-00193788.
XX

```

```

PR 24-JUN-1998; 98JP-00193788.
XX

```

```

PA (SEKG ) SEIKAGAKU KOGYO CO LTD.
XX

```

```

XX WPI; 2000-140125/13.
XX

```

```

XX A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.
XX

```

```

XX Claim 4; Page; 30pp; Japanese.
XX

```

```

XX The present invention describes a recombinant protein which consists of
CC three continuous regions (N-terminal region, internal region and C-
CC terminal region) where one or two regions among the above three regions
CC is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. AAV78127 to AAV78142 represent specifically claimed
CC examples of recombinant proteins from the present invention. N.B. The
CC present sequence is not given in the present specification, but is
CC derived from sequences given as specified in the claim
XX

```

```

SQ Sequence 563 AA;

```

```

Query March 66.4%; Score 1238; DB 3; Length 563;
Best Local Similarity 80.8%; Pred. No. 2.9e-129;
Matches 252; Conservative 10; Mismatches 28; Indels 22; Gaps 6;

```

```

QY 21 RRVLTIAPALLILGLMTWYAAGVPLASDRYGLLAFLGYGAFSAHLVAQSLFAYLEHRR 79
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
11 RIIGTTTFGVSLLLGIIITAAVIVGQFIQTDNY-YFSFGLYGA-----LFAYLEHRR 59
QY 80 RVAAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCILASARALYPPARLRVLMV 134
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
60 RVAAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCILTSARALYPPHTRLRVLMV 119
QY 135 DGNRAEDLYVDMFREVFADEDPATYVWDGNYHQWPEPA-AAGAVGAGAYREVEAEDPGR 193
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
120 DGNRAEDLYVDMFREVFADEDPATYVWDGNYHQWPEPAEATGAVGEGAYREVEAEDPGR 179
QY 194 LAVEALVTRRCVCVQAORWGKREVMYTAFKALGDSVDYVQVCSDTRLDPMALLELRV 253
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
180 LAVEALVTRRCVCVQAORWGKREVMYTAFKALGDSVDYVQVCSDTRLDPMALLELRV 239
QY 254 LDEDPRVGAAGVDVRLNPLDSWVSFLSSLRVWVAFNVERACQSYFHCVCISGPLESCP 313
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
240 LDEDPRVGAAGVDVRLNPLDSWVSFLSSLRVWVAFNVERACQSYFHCVCISGPLE---- 295

```

```

QY 314 GPREHAMMPSFL 325
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
296 GLYRNNLLQQFL 307

```

RESULT 14

```

AAV78128
ID AAV78128 standard; protein; 577 AA.
XX
AC AAV78128;
XX

```

```

DT 27-APR-2000 (first entry)
XX

```

```

DE Recombinant chimeric hyaluronate synthase modified protein #2.
XX

```

```

KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX

```

```

OS Mus sp.
XX

```

```

OS Synthetic.
XX

```

```

OS Chimeric.
XX

```

```

FN JP2000004886-A.
XX

```

```

PD 11-JAN-2000.
XX

```

```

PF 24-JUN-1998; 98JP-00193788.
XX

```

```

PR 24-JUN-1998; 98JP-00193788.
XX

```

```

PA (SEKG ) SEIKAGAKU KOGYO CO LTD.
XX

```

```

XX WPI; 2000-140125/13.
XX

```

```

XX A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.
XX

```

```

XX Claim 3; Page; 30pp; Japanese.
XX

```

```

XX The present invention describes a recombinant protein which consists of
CC three continuous regions (N-terminal region, internal region and C-
CC terminal region) where one or two regions among the above three regions
CC is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
XX

```


THIS PAGE BLANK (138710)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 14:06:48 ; Search time 21.1364 Seconds
(without alignments)
1638.789 Million cell updates/sec

Title: US-10-672-399-6

Perfect score: 1864

Sequence: 1 MRQDAPKPTPAARCSGLA.....ALVLPAAVEADTLVQVVLP 360

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_79:*

1: Pirl:*

2: Pirl:*

3: Pirl:*

4: Pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1429	76.7	543	2 JC4812	hyaluronan synthas
2	804.5	43.2	588	2 A43740	DG42 protein - Afr
3	310	16.6	568	2 T17588	hyaluronoglucosami
4	214	11.5	424	1 ZZRCL	modulation protein
5	204.5	11.0	426	2 A95321	NodC N-ACETYLGLUCO
6	201.5	10.8	426	1 ZZRQ4	modulation protein
7	194.5	10.4	413	1 S34305	modulation protein
8	173.5	9.3	424	1 S12793	modulation protein
9	163.5	8.8	1086	2 JC6079	chitin synthase (E
10	159	8.5	1195	2 S61886	chitin synthase (E
11	158	8.5	395	1 JQ0396	modulation protein
12	156.5	8.4	1165	1 S45879	chitin synthase (E
13	151.5	8.1	419	2 A53100	hyaluronate synthas
14	151.5	8.1	1112	2 T30202	probable chitin sy
15	149.5	8.0	395	2 A48755	hyaluronan synthas
16	149.5	8.0	1239	2 T42020	class IV chitin sy
17	145.5	7.8	1041	2 T31037	chitin synthase (E
18	142	7.6	1175	2 S39951	chitin synthase (E
19	137.5	7.4	365	2 E59102	hypothetical prote
20	135	7.2	428	1 E38180	modulation protein
21	123.5	6.6	1103	2 T42022	probable chitin sy
22	117	6.3	743	2 T34632	probable bi-functi
23	116	6.2	447	2 A97211	glycosyltransferas
24	114.5	6.1	1498	2 S78102	chitin synthase (E
25	114.5	6.1	1852	2 JC5546	chitin synthase (E
26	114.5	6.1	1869	2 A59290	class V chitin syn
27	111.5	6.0	412	2 B90075	intercellular adhe
28	108.5	5.8	869	2 H83500	probable glucosyl
29	105	5.6	615	2 E70663	probable PPE prote

30 104.5 5.6 3519 2 S43048 polyketide synthas
31 101.5 5.4 1198 2 T28678 polyketide synthas
32 100.5 5.4 3034 2 T14119 seven-pass transme
33 100 5.4 1155 2 AC2675 chromosome segrega
34 100 5.4 1165 2 A57457 structural mainten
35 99.5 5.3 291 2 B82645 phenylacetalddehyde
36 99.5 5.3 412 2 S77608 probable intercell
37 99 5.3 352 2 F75099 rhannosyl transfer
38 98.5 5.3 270 2 A82313 probable beta-keto
39 97 5.2 414 2 S18962 FB15 protein - St
40 97 5.2 482 2 G83928 hypothetical proce
41 97 5.2 871 2 AE3085 two component sens
42 97 5.2 880 2 D98201 hypothetical prote
43 97 5.2 3573 2 S23070 erythronolide synt
44 96.5 5.2 420 2 D69769 cellulose synthase
45 96.5 5.2 460 2 H87396 conserved hypothet

ALIGNMENTS

RESULT 1

JC4812

hyaluronan synthase (EC 2.4.1.1-) - human

C;Species: Homo sapiens (man)

C;Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 02-Mar-2001

C;Accession: JC4812

R;Itano, N.; Kimata, K.

Biochem. Biophys. Res. Commun. 222, 816-820, 1996

A;Title: Molecular cloning of human hyaluronan synthase.

A;Reference number: JC4812; MUID:96244584; PMID:8651928

A;Accession: JC4812

A;Molecule type: mRNA

A;Residues: 1-543 <I>A

A;Cross-references: DDBJ:D84424; NID:G1401033; PIDN:BAAL12351.1; PID:d1013030; PID:g14010

C;Comment: This enzyme synthesizes a high molecular weight glycosaminoglycan composed of

C;Genetics:

A;Gene: GDB:HAS1; HAS

A;Cross-references: GDB:1220109; OMIM:601463

A;Map position: 19q13.4-19q13.4

C;Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein

F;58,167,308,321/Binding site: phosphate (Thr) (covalent) #status predicted

F;82,247/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 76.7%; Score 1429; DB 2; Length 543;
Best Local Similarity 95.2%; Pred. No. 4.8e-114; Mismatches 6; Indels 4; Gaps 1;
Matches 276; Conservative 4;

Qy 36 MTWAYAGVPLASDRYGLLAFLGYSALHVAQSLFAYLEHRRVAAAARGPLDAATAR 95

Db 1 MTWAYAGVPLASDRYGLLAFLGYSALHVAQSLFAYLEHRRVAAAARGPLDAATAR 60

Qy 96 SVALTISAYQEDPAYLRQCCLASARALLYPRARLVLMVVDGNRAEDLVMDMFREVFADE 155

Db 61 SVALTISAYQEDPAYLRQCCLASARALLYPRARLVLMVVDGNRAEDLVMDMFREVFADE 120

Qy 156 DPATVYWDGNYHQWEPAPAGAGVAGAYREVEDPGRGLAVEALVTRRCVCVAQRWGK 215

Db 121 DPATVYWDGNYHQWEPAPAGAGVAGAYREVEDPGRGLAVEALVTRRCVCVAQRWGK 180

Qy 216 REVMTYAFKALGSDVDYVQVCDSTRLDPMALLBLRVLDPRVGVAGGVDVRLINPLDS 275

Db 181 REVMTYAFKALGSDVDYVQVCDSTRLDPMALLBLRVLDPRVGVAGGVDVRLINPLDS 240

Qy 276 WVSFSLSLRYVWAFNVERACQSYFHCVSCISGPLESCFPGPREHAMPSFL 325

Db 241 WVSFSLSLRYVWAFNVERACQSYFHCVSCISGPLESCFPGPREHAMPSFL 286

RESULT 2

A43740

DG42 protein - African clawed frog

C;Species: Xenopus laevis (African clawed frog)

A:Reference number: JQ0393; MUID:90136519; PMID:2615763
A:Accession: JQ0396
A:Molecule type: DNA
A:Residues: 1-395 <GOE>
A:CROSS-references: UNIPROT:Q07755; GB:L18897; NID:gl293899; PIDN:AAB51164.1; PID:g31029
A:Experimental source: strain ORS571
C:Comment: this is one of the proteins, coded by nodulation genes, that are required for
C:Genetics:
A:Gene: nodC
C:Superfamily: nodulation protein nodC
C:Keywords: nodulation

Query Match 8.5%; Score 158; DB 1; Length 395;
Best Local Similarity 26.1%; Pred. No. 9.4e-06;
Matches 87; Conservative 36; Mismatches 112; Indels 98; Gaps 16;

QY 49 DRYGLAFLGLYCAFLSAHLVAQSLFAYLEHREVAARGLDA--ATARSVALTISAYOE 106
DB 5 DVIGLLATAAYVTLASAYKVQ-----FINVSVTDVAGLESDALPLTPR-VDVIVPTFNE 59
QY 107 DPAYLRQCILASARALLYPRARLRVLMVVDGNRAE-----DLYMVDMFREVFADEDPAT 159
DB 60 NSTLLECVASICAQDY-RGPTIVVDDGSTNKTSPHAVCDKYASDE-RFIFVELD--- 114
QY 160 VYMDGNHYQHPWEPAAGAVGAGAYREVEABDPGRILA-VEALVTRRCVCVCAQRWGGKREV 218
DB 115 -----QNKGTAAQMEAIRRT----- 129
QY 219 MYTAFKALGDSVDYVQVCDSTRDLPMLLELRVLDDEPRVAGVGVDVRLNPLDSWVS 278
DB 130 -----DGLLILNVDSDVIDKVVTKLASSM-RAPNVGVMGQVLVAKNRSMILT 178
QY 279 FLSSLRVYVAFNVERACOSYHCVSCISGPLSCPGPREHAMMP-----SELA-PVQ 329
DB 179 RLIDNEYWLACNEERIAQSFQSVWCCCGP---CAMYRSALTPLLAEYEHOTFLGRPSN 235
QY 330 V-----HLQVPLL---LRDALVLPVAVEPADTIV 355
DB 236 FGEDRHLITLMKAGFRTGYVPSAVAR---TLV 265

RESULT 12
S45879
chitin synthase (EC 2.4.1.16) 3 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YBR023c; protein YBR0305
C:Species: Saccharomyces cerevisiae
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: S45879; S46554; S22776; A39639; S17247
R:Grivell, L.A.; de Haan, M.; Maat, M.J.; Smits, P.H.M.
submitted to the Protein Sequence Database, August 1994
A:Reference number: S45875
A:Accession: S45879
A:Molecule type: DNA
A:Residues: 1-1165 <GRI>
A:CROSS-references: UNIPROT:P29465; EMBL:235892; NID:g536229; PIDN:CAA84965.1; PID:g5362
A:Experimental source: strain S288C
R:Smits, P.H.M.; de Haan, M.; Maat, C.; Grivell, L.A.
Yeast 10 (Suppl.A), S75-S80, 1994
A:Title: The complete sequence of a 33 kb fragment on the right arm of chromosome II fro
ly identified genes and a homologue of the SCO1 gene.
A:Reference number: S46551; MUID:94378725; PMID:8091864
A:Accession: S46554
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1165 <SMI>
A:CROSS-references: EMBL:X76078; NID:g498748; PIDN:CAA53680.1; PID:g498752
A:Experimental source: strain S288C
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993
R:Bulawa, C.E.
Mol. Cell. Biol. 12, 1764-1776, 1992
A:Title: CSD2, CSD3, and CSD4, genes required for chitin synthesis in Saccharomyces cere
blum species and xenopus laevis.
A:Reference number: S22776; MUID:92195323; PMID:1532231

A:Accession: S22776
A:Molecule type: DNA
A:Residues: 1-1162, 'L', 1164-1165 <BUL>
A:CROSS-references: EMBL:M73697; NID:gl72103; PIDN:AAA34844.1; PID:gl72104
R:Valdivieso, M.H.; Mol, P.C.; Shaw, J.A.; Cabib, E.; Duran, A.
J. Cell Biol. 114, 101-109, 1991
A:Title: CAL1, a gene required for activity of chitin synthase 3 in Saccharomyces cerevi
A:Reference number: A39639; MUID:91268144; PMID:2050737
A:Accession: A39639
A:Molecule type: DNA
A:Residues: 67-1165 <VAL>
A:CROSS-references: GB:X57300; NID:g3359; PIDN:CAA40559.1; PID:g3360
C:Genetics:
A:Gene: SGD:CHS3; CAL1; CSD2; MIPS:YBR023c
A:CROSS-references: MIPS:YBR023c; SGD:S0000227
A:Map position: 2R
C:Function:
A:Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucosa
C:Superfamily: chitin synthase chs4
C:Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase; transmembrane
F:203-219/Domain: transmembrane #status predicted <TM1>
F:457-473/Domain: transmembrane #status predicted <TM2>
F:1018-1034/Domain: transmembrane #status predicted <TM3>
F:1035-1054/Domain: transmembrane #status predicted <TM4>
F:1060-1076/Domain: transmembrane #status predicted <TM5>
F:1084-1100/Domain: transmembrane #status predicted <TM6>

Query Match 8.4%; Score 156.5; DB 1; Length 1165;
Best Local Similarity 22.2%; Pred. No. 4.5e-05;
Matches 58; Conservative 48; Mismatches 88; Indels 67; Gaps 9;

QY 101 ISAVQEDPAYLRQCILASARALLYPRARLRVLMVVDG-----NRAED-----LYMVDMFRE 150
DB 658 VTCYSEDEGLRTLDLSLTDDYPSNSHKLVLVVCGLKSGNDKTTPEIALGWMDDFTV 717
QY 151 VPAEDPATYVWDGNHYQHPWEPAAGA-----VGAGAYR-----EVAEDPGRILAVAL 199
DB 718 PPDEVKPYSYV-----AVASGSKRHNAKIYAGFYKYDDSTIPPENQQRVPPIITI 767
QY 200 VTRRCVCVAQRWG-----GKR--EVMTATF-----KAL 226
DB 768 V---KCGTPAEQGAAPGNRGRDSQIILMSFLEKITFDERTWOLFOLLKNWIQTGLM 824
QY 227 GDSVDYVQVCDSTRDLPMLLELRVLDDEPRVAGVGVDVRLNPLDSWVSFLSSLRVW 286
DB 825 ADFVETVLVMDADIKVFPDALTHVAVEMVKDPLINGLGLGETKIANKAQSWTAIQVFYY 884
QY 287 VAFNVERACOSYHCVSCISG 307
DB 885 ISHQAKAFESVFGSVTCCLPG 905

RESULT 13
A53100
hyaluronate synthase A (HaseA) - Streptococcus sp. (group A)
C:Species: Streptococcus sp.
C>Date: 06-Oct-1994 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: A53100
R:Dougherty, B.A.; van de Rijn, I.
J. Biol. Chem. 269, 169-175, 1994
A:Title: Molecular characterization of haseA from an operon required for hyaluronic acid
A:Reference number: A53100; MUID:94103204; PMID:8267791
A:Accession: A53100
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-419 <DOU>
A:CROSS-references: UNIPROT:Q8NKK1
A:Note: sequence extracted from NCBI backbone (NCBIN:141683, NCBI:P:141684)
C:Superfamily: nodulation protein nodC

Query Match 8.1%; Score 151.5; DB 2; Length 419;
Best Local Similarity 24.0%; Pred. No. 3.6e-05;
Matches 80; Conservative 38; Mismatches 116; Indels 99; Gaps 13;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 13:48:17 ; Search time 94.0909 Seconds
(without alignments)
1959.259 Million cell updates/sec

Title: US-10-672-399-6

Perfect score: 1864

Sequence: 1 MRQDAPKPTPAARRCSGLA.....ALVLPVAVAPADTLVQVVLP 360

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1605	86.1	578	1 HAS1 HUMAN	Q92839 homo sapien
2	1597	85.7	578	2 Q9NS49	Q9ns49 homo sapien
3	1588	85.2	577	2 Q81YH3	Q81yh3 homo sapien
4	1545.5	82.9	583	2 Q6S742	Q6s742 papio anubi
5	1512	81.1	583	1 HAS1 MOUSE	Q61647 mus musculus
6	1507	80.8	583	2 Q8CH93	Q8ch93 rattus norv
7	929.5	49.9	458	2 Q6T488	Q6t488 brachydanio
8	811	43.5	552	2 Q95M29	Q95m29 oryctolagus
9	808	43.3	557	2 Q6W9J2	Q6w9j2 xenopus lae
10	804.5	43.2	588	1 HAS1 XENLA	P13563 xenopus lae
11	803.5	43.1	554	2 Q9DG30	Q9dg40 brachydanio
12	800.5	42.9	553	2 Q75R37	Q75r37 sus scrofa
13	800	42.8	554	1 HAS3 MOUSE	Q8650 mus musculus
14	798	42.8	554	2 Q8CEB9	Q8ceb9 mus musculus
15	798	42.8	554	2 Q8CH92	Q8ch92 rattus norv
16	790.5	42.4	553	1 HAS3 HUMAN	O00219 homo sapien
17	790.5	42.4	553	2 Q96R72	Q96r72 homo sapien
18	784.5	42.1	552	2 Q95M16	Q95m16 oryctolagus
19	783.5	42.0	552	1 HAS2 HUMAN	Q92819 homo sapien
20	783.5	42.0	552	1 HAS2 MOUSE	P70312 mus musculus
21	783.5	42.0	552	2 Q8H2J3	Q8hj3 equus caball
22	782.5	42.0	552	2 Q8S070	Q8sq70 sus scrofa
23	781.5	41.9	552	1 HAS2 BOVIN	Q9711 bos taurus
24	780	41.8	552	2 Q9DG41	Q9dg41 brachydanio
25	778.5	41.8	552	1 HAS2 RAT	Q35776 rattus norv
26	771	41.4	552	1 HAS2 CHICK	O57424 gallus gall
27	750.5	40.3	551	1 HAS2 XENLA	O57427 xenopus lae
28	739.5	39.7	583	1 HAS2 XENLA	O57428 xenopus lae
29	738.5	39.6	583	2 Q6AZT0	Q6azt0 xenopus lae
30	592.5	31.8	281	2 Q8WT20	Q8wt20 homo sapien
31	560.5	30.1	131	2 Q8BPN0	Q8bpn0 mus musculus

32 541 29.0 393 2 Q811Y6 Q811y6 rattus norv
33 351 18.8 156 2 O18792 O18792 papio anubi
34 310 16.6 568 2 Q84419 Q84419 paramescium
35 310 16.6 568 2 Q9WFS9 Q9wfs9 paramescium
36 310 16.6 568 2 Q9WFT0 Q9wft0 paramescium
37 291 15.6 134 2 Q90489 Q90489 brachydanio
38 272 14.6 245 2 Q9GK14 Q9gk14 bos taurus
39 214.5 11.5 452 2 Q9AQ23 Q9aq23 bradyrhizob
40 214 11.5 424 1 NODC RHILV P04340 rhizobium l
41 204.5 11.0 426 1 NODC RHIME P04341 rhizobium m
42 201 10.8 452 2 Q6EX51 Q6ex51 sinorhizobi
43 198.5 10.6 408 2 Q6PTX8 Q6ptx8 rhizobium s
44 196.5 10.5 413 1 NODC RHISN P50357 rhizobium s
45 195 10.5 443 2 Q8KLG3 Q8klg3 rhizobium e

ALIGNMENTS

RESULT 1

HAS1_HUMAN STANDARD; PRT; 578 AA.
AC Q92839; Q14470;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE (Hyaluronic acid synthase 1) (HA synthase 1) (HuHAS1).
GN Name=HAS1; Synonyms=HAS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph node;
RX MEDLINE=96394438; PubMed=8798544; DOI=10.1074/jbc.271.38.23395;
RA Shyjan A.M., Heldin P., Butcher E.C., Yoshino T., Brieskin M.J.;
RT "Functional cloning of the cDNA for a human hyaluronan synthase.";
RL J. Biol. Chem. 271:23395-23399(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=96244584; PubMed=8651928; DOI=10.1006/bbrc.1996.0827;
RA Itano N., Kimata K.;
RT "Molecular cloning of human hyaluronan synthase.";
RL Biochem. Biophys. Res. Commun. 222:816-820(1996).
CC -!- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA) synthesis. Also able to catalyze the synthesis of chito-oligosaccharide depending on the substrate.
CC -!- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-glucuronosyl(1->3)](n) + 2n UDP.
CC -!- COFACTOR: Magnesium.
CC -!- PATHWAY: Hyaluronate synthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: Highly expressed in ovary followed by spleen, thymus, prostate, testes and large intestine. Weakly expressed in small intestine.
CC -!- SIMILARITY: Belongs to the nodC/HAS family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U59269; AAC50706.1; --
EMBL; D84424; BAA12351.1; ALT_INIT.
Genew; HGNC:4819; HAS1.
MIM; 601463; --

RL	Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AY463695; AAR25554.1; ..
DR	InterPro; IPR001173; Glyco trans 2.
DR	InterPro; IPR000276; GPCR Rhodspan.
DR	InterPro; IPR002057; Isopen_N synth.
DR	PROSITE; PS00237; G_PROTEIN_RECPE_F1_1; UNKNOWN_1.
DR	PROSITE; PS00185; IPNS_1; UNKNOWN_1.
SQ	SEQUENCE 583 AA; 65596 MW; EA47354C89346F94 CRC64;
Query Match 82.9%; Score 1545.5; DB 2; Length 583;	
Best Local Similarity 91.5%; Pred. No. 4e-121;	
Matches 302; Conservative 6; Mismatches 13; Indels 9; Gaps 2	
Qy	1 MRQADAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFLGLYG 60
Db	1 MTQRTDKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFLGLYG 60
Qy	61 AFI SAHLVAQSLFAYLEHRRVA-----AARGPLDAATARSVALTISAYQEDPAYLRQCL 115
Db	61 AFI SAHLVAQSLFAYLEHRRVAARAAARAAAGRLDAATARSVALTISAYQEDPAYLRQCL 120
Qy	116 ASARALLYPARLRLVLMVDPGNRAEDLYVDMDFEVFADEDPAIYVWDGNYHQWEPAAA 175
Db	121 VSARALLYPARLRLVLMVDPGNRPEDLYVDMDFEVFADEDPAIYVWDGNYHQWEPAAV 180
Qy	176 GAVGAGAYREVEAEDPGRLAVEALVTRTRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQV 235
Db	181 GAVGVGAYREVEAEDPGRLAVEALVTRTRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQV 240
Qy	236 CDS DTRLDPMALLELVRLVLEDDPRVGA VGDVRLINPLDSDVSVFLSSLYRWVAFNVERAC 295
Db	241 CDS DTRLDPMALLELVQVLEDDPRVGA VGDVRLINPLDSDVSVFLSSLYRWVAFNVERAC 300
Qy	296 QSYFHCVCISCGPLESCPGPREHAMPSFL 325
Db	301 QSYFHCVCISGGLP-----GLYRNLLQQFL 326
RESULT 5	
HA	HAS1_MOUSE STANDARD; PRT; 583 AA.
ID	Q61647;
AC	30-MAY-2000 (Rel. 39, Created)
DT	DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE	05-JUL-2004 (Rel. 44, Last annotation update)
DE	Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE	(Hyaluronic acid synthase 1) (HA synthase 1).
GN	Name=Has1; Synonyms=Has;
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RP	[1]
RP	SEQUENCE FROM N.A.
RP	MEDLINE=96215261; PubMed=8626618; DOI=10.1074/jbc.271.17.9875;
RA	Itano N.; Kimata K.;
RT	"Expression cloning and molecular characterization of HAS protein, a
RT	eukaryotic hyaluronan synthase.";
RT	J. Biol. Chem. 271:9875-9878(1996).
RP	[2]
RP	MUTAGENESIS.
RA	MEDLINE=20085071; PubMed=10617644; DOI=10.1074/jbc.275.1.497;
RA	Yoshida M.; Itano N.; Yamada Y.; Kimata K.;
RT	"In vitro synthesis of hyaluronan by a single protein derived from
RT	mouse HAS1 gene and characterization of amino acid residues essential
RT	for the activity.";
RT	J. Biol. Chem. 275:497-506(2000).
CC	-I- FUNCTION: plays a role in hyaluronan/hyaluronic acid (HA)
CC	synthesis. Also able to catalyze the synthesis of chito-
CC	oligosaccharide depending on the substrate.
CC	-I- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-
CC	glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-
CC	glucuronosyl(1->3)](n) + 2n UDP.

RESULT 9

ID Q6W9J2 PRELIMINARY; PRT; 557 AA.
 AC Q6W9J2;
 DT 05-JUL-2004 (TRMBLrel. 27, Created)
 DT 05-JUL-2004 (TRMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRMBLrel. 27, Last annotation update)
 DE Hyaluronic acid synthase 3.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OC NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14667843; DOI=10.1016/S0945-053X(03)00082-9;
 RA Vignetti D., Viola M., Gornati R., Ori M., Passi A.,
 RA De Luca G., Bernardini G.
 RT "Molecular cloning, genomic organization and development expression
 of the Xenopus laevis hyaluronan synthase 3.";
 RL Matrix Biol. 22:511-517(2003).
 DR EMBL; AY302252; AAP58398.1; --
 DR InterPro; IPR001173; Glyco_transf.2.
 DR Pfam; PF00535; Glycos.transf.2; 1.
 SQ SEQUENCE 557 AA; 64060 MW; F81478B685BF7AD5 CRC64;

[illegible]

Db	164	QVBMVKNFQVVCIMQKGGKREVYTAFRALGDSVAVVQVCDSDTVLPACTAEMLRILE	223
Qy	256	EDPRVGAVGGDVRIINPLDWSVLSLSIRYVVAFNVERACOSYFHCVSCISPLESCPGP	315
Db	224	EDSEVGGVGGDVQILNKYENWISFLSPRYMAFNVERACOSYFCVCISGPL----	Qm 279

```

QY 316 REHAMPSPFL 325
Db 280 YRNSLLQYFL 289

RESULT 10
HAS1_XENLA
ID HAS1_XENLA STANDARD; PRT; 588 AA.
AC P13563;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE (Hyaluronic acid synthase 1) (HA synthase 1) (XHASI) (DG42 protein).
GN Name=HAS1;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88313363; PubMed=3410156;
RA Rosa F., Sargent T.D., Rebbert M.L., Michaels G.S., Jamrich M.,
RA Grunz H., Jonas E., Winkler J.A., Dawid I.B.;
RT "Accumulation and decay of DG42 gene products follow a gradient
RT pattern during Xenopus embryogenesis.";
RL Dev. Biol. 129:114-123(1988).
CC -!- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA)
CC synthesis (By similarity). May play a role in signaling or pattern
CC formation in embryonic development.
CC -!- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-
CC glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-
CC glucuronosyl(1->3)](n) + 2n UDP.
CC -!- COFACTOR: Magnesium.
CC -!- PATHWAY: Hyaluronate synthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: Embryo.
CC -!- DEVELOPMENTAL STAGE: During gastrula and early neurula stages.
CC -!- SIMILARITY: Belongs to the nodC/HAS family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22249; AAA49699.1; -;
DR PIR; A43740; A43740.
DR InterPro; IPR001173; Glyco trans 2.
KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
FT DOMAIN 1 28
FT TRANSMEM 29 49
FT DOMAIN 50 61
FT TRANSMEM 62 82
FT DOMAIN 83 411
FT TRANSMEM 412 432
FT DOMAIN 433 433
FT TRANSMEM 434 454
FT DOMAIN 455 456
FT TRANSMEM 457 477
FT DOMAIN 478 505
FT TRANSMEM 506 526
FT DOMAIN 527 543
FT TRANSMEM 544 564
FT DOMAIN 565 588
FT TRANSMEM 588 AA; 68522 MW; 33DA3B8E331F4CE9 CRC64;
SQ SEQUENCE 588 AA; 68522 MW; 33DA3B8E331F4CE9 CRC64;

Query Match 43.2%; Score 804.5; DB 1; Length 588;
Best Local Similarity 53.6%; Pred. No. 6.8e-59;

DR EMBL; M22249; AAA49699.1; -;
DR PIR; A43740; A43740.
DR InterPro; IPR001173; Glyco trans 2.
KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
FT DOMAIN 1 28
FT TRANSMEM 29 49
FT DOMAIN 50 61
FT TRANSMEM 62 82
FT DOMAIN 83 411
FT TRANSMEM 412 432
FT DOMAIN 433 433
FT TRANSMEM 434 454
FT DOMAIN 455 456
FT TRANSMEM 457 477
FT DOMAIN 478 505
FT TRANSMEM 506 526
FT DOMAIN 527 543
FT TRANSMEM 544 564
FT DOMAIN 565 588
FT TRANSMEM 588 AA; 68522 MW; 33DA3B8E331F4CE9 CRC64;
SQ SEQUENCE 588 AA; 68522 MW; 33DA3B8E331F4CE9 CRC64;

Query Match 43.2%; Score 804.5; DB 1; Length 588;
Best Local Similarity 53.6%; Pred. No. 6.8e-59;

Matches 156; Conservative 51; Mismatches 75; Indels 9; Gaps 3;
QY 22 RVLTIAFALLILGLMTWAYAAGVPLASDRYCLLAFGLYGAFLSAHLVAQSLFAYLEHRRV 81
Db 28 RIITYSEGVLLATITAAAYVAEFQVLKHEATLSGLYGLAMLLHLMQSLFAFLIEIRV 87
QY 82 AAAARGPLDAATARSVALTISAYOEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAED 141
Db 88 ---NKSELPSCFKKTVALTITAGYOENPEYLKCLSECKYKPKDKLKIILVDGNTDEDD 144
QY 142 LYMDMFREVPFADDPATYVMDGNHYHQFWEPAAGAGAGAGAYREVEA---EDPGRLAYEA 198
Db 145 AYMMEMKDFVHGSDVGTYYWKNYHTVKKPE---ETNKGSCPEVSKPLNDEDEGINMVEE 201
QY 199 LVTRRCVCVAQRWGGKREVMYTAFAKLGSDVDYVQVCDSDTRLDPMALLBLRVLDSDP 258
Db 202 LVNRKRCVCIMQWGGKREVMYTAFAIGTSVDYVQVCDSDTKLDELATVEMVKVLESND 261
QY 259 RVGAVGGDVRILNPLDSWVSFLSSRLRYWVAENVERACQSYFHCVSCISGGL 309
Db 262 MYGAVGGDVRILNPLDYSFISFMSSRLRYWMAFNVERACQSYFDCVSCISGGL 312

RESULT 11
Q9DG40
ID Q9DG40 PRELIMINARY; PRT; 554 AA.
AC Q9DG40;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hyaluronan synthase 3.
GN Name=has3;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14729574;
RA Bakkens J., Kramer C., Pothof J., Quaedvlieg N.E., Spaik H.P.,
RA Hammerschmidt M.;
RT "Has2 is required upstream of Rac1 to govern dorsal migration of
RT lateral cells during zebrafish gastrulation.";
RL Development 131:525-537(2004).
DR EMBL; AF190743; AAG17181.1; -;
DR ZFIN; ZDB-GENE-021118-1; has3.
DR InterPro; IPR001173; Glyco trans 2.
SQ SEQUENCE 554 AA; 63077 MW; 26273CB459CE3E97 CRC64;

Query Match 43.1%; Score 803.5; DB 2; Length 554;
Best Local Similarity 54.0%; Pred. No. 7.7e-59;
Matches 169; Conservative 44; Mismatches 79; Indels 21; Gaps 9;
QY 18 GLARRV-LTATFALLILGLMTWAYAAGVPLASDRYCLLAFGLYGAFLSAHLVAQSLPAYL 76
Db 6 GTAVRIFITTLFAAVLVFAILLALVVTGYQFIHTQHLLSPGLYGAFLSHLLQLSLPAYL 65
QY 77 EHRVAAAARGP-LDAATARSVALTISAYQDDPAYLRQCLASARALLYPRARLRVLMVVD 135
Db 66 EHRQW---RGPSRPQHLRRTVALCIAAYQDDPAYLRKCLSSR-ISFP--GLKVLVVD 118
QY 136 GNRADLYMDMFREVPFAD-EDPATYVMDGNHYHQFWEPAAGAGAGAGAYREVEADPRL 194
Db 119 GNRQEDAYMMDIFQEVNMGVGEQTGCVVMKGNHNSGDDGGGGGKGS-----VHAEEARV 173
QY 195 AVEALVTRRCVCVAQRWGGKREVMYTAFAKLGSDVDYVQVCDSDTRLDPMALLBLRVLD 254
Db 174 A--RVRSRCVSCIMQWGGKREVMYTAFAKLGSDVDYVQVCDSDTVLDPMCTEMKIL 231
QY 255 DEDPRVAVGGDVRILNPLDSWVSFLSSRLRYWVAENVERACQSYFHCVSCISGPLESCPG 314
Db 232 EEDPDVGGVGGDVQILNKYDSWISFLSSRVYMAFNVERACQSYFGCVQCISGPL----G 287

```


QY 315 PREHAMPSFLAP 327
Db 288 MYNSLLQOFLP 300

RESULT 12

Q75R37 PRELIMINARY; PRT; 553 AA.
AC Q75R37;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hyaluronan synthase 3.
GN Name=shas3;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9833;
RN [1]
RP SEQUENCE FROM N.A.
RA Kimura N., Konno Y., Yokoo M., Sato E.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB159675; BAD1881.1; -;
DR InterPro; IPR001173; Glyco_trans_2.
SQ SEQUENCE 553 AA; 63110 MW; C97EBE911723B44C CRC64;

Query Match 42.9%; Score 800.5; DB 2; Length 553;
Best Local Similarity 54.1%; Pred. No. 1.4e-58;
Matches 166; Conservative 38; Mismatches 84; Indels 19; Gaps 5;
QY 21 RRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAGLYGAFLSAHLVAQSLFAYLEHRR 80
Db 10 RVVGTSLFALVGLGILAAVYTGQFIHTEKHVLSFGLYGAILGLHLILQSLFAFLEHRR 69
QY 81 VAAARG-PLDAAATRSVALTISAVQEDPAYLRQCLASARALLYPPARLRVLMVDGNRA 139
Db 70 MRRAGRPLKPLQRSRVALCIAAYQEDPDYLRKCLRSQAQRIAFP--DLKVVVMVDGNRQ 127
QY 140 EDLYVMDVMEFVF-ADEDPATYVWDGNVHQPPEAAAGVAGAGAYREVEAEDPGRALVEA 198
Db 128 EDAYMLDIFHEVLGGTEQAGFFVWRNSFNFEAGEGETEASLQEGMR-----VRN 176
QY 199 LVTRRCVCVAQRWGGKREVMYTAFAKALGSDVDYVQVCDSDTRLDPMALLELRVLDDEP 258
Db 177 VVRASTFCIMQKMGKREVMYTAFAKALGSDVDYVQVCDSDTVLDPACTFEMLRVLEEDP 236
QY 259 RVGAVGGDVRIINPLDSWVSFLSSLRVYVAFNVERACQSYFHCVCISGPLESCPGPREH 318
Db 237 QVGGVGGDVQILNKYDSWISFLSSVRYVWAFNVERACQSYFHCVCISGPLESCPGPREH 318
QY 319 AMMPSPFL 325
Db 293 SLLQOFL 299

RESULT 13

HAS3 MOUSE
ID HAS3 MOUSE STANDARD; PRT; 554 AA.
AC O08650;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hyaluronan synthase 3 (FC 2.4.1.212) (Hyaluronate synthase 3)
DE Hyaluronic acid synthase 3) (HA synthase 3).
GN Name=Has3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ, and C57BL/6;

RX MEDLINE=97238820; PubMed=9083017; DOI=10.1074/jbc.272.14.8957;
RA Spicer A.P., Olson J.S., McDonald J.A.;
RT "Molecular cloning and characterization of a cDNA encoding the third
RT putative mammalian hyaluronan synthase.";
RL J. Biol. Chem. 272:8957-8961(1997).
RC -!- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA)
CC synthesis.
CC -!- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-
CC glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-
CC glucuronosyl(1->3)](n) + 2n UDP.
CC -!- COFACTOR: Magnesium.
CC -!- PATHWAY: Hyaluronate synthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- DEVELOPMENTAL STAGE: Expressed at E17.5 day.
CC -!- SIMILARITY: Belongs to the nodC/HAS family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; U86408; AAC53128.1; -;
DR MGD; MGI:109599; Hae3.
DR InterPro; IPR001173; Glyco_trans_2.
KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
FT DOMAIN 1 15 Cytoplasmic (Potential).
FT TRANSMEM 16 36 Extracellular (Potential).
FT DOMAIN 37 44 Extracellular (Potential).
FT TRANSMEM 45 65 Cytoplasmic (Potential).
FT DOMAIN 66 378 Cytoplasmic (Potential).
FT TRANSMEM 379 399 Extracellular (Potential).
FT DOMAIN 400 409 Extracellular (Potential).
FT TRANSMEM 410 430 Cytoplasmic (Potential).
FT DOMAIN 431 441 Cytoplasmic (Potential).
FT TRANSMEM 442 462 Extracellular (Potential).
FT DOMAIN 463 474 Extracellular (Potential).
FT TRANSMEM 475 495 Cytoplasmic (Potential).
FT DOMAIN 496 516 Cytoplasmic (Potential).
FT TRANSMEM 517 537 Extracellular (Potential).
FT DOMAIN 538 554 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 463 463
SQ SEQUENCE 554 AA; 63338 MW; 88296DCB465CE3EC CRC64;

Query Match 42.9%; Score 800; DB 1; Length 554;
Best Local Similarity 53.7%; Pred. No. 1.5e-58;
Matches 166; Conservative 41; Mismatches 80; Indels 22; Gaps 6;

QY 21 RRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAGLYGAFLSAHLVAQSLFAYLEHRR 80
Db 10 RVVGTSLFALVGLGILAAVYTGQFIHTEKHVLSFGLYGAILGLHLILQSLFAFLEHRR 69
QY 81 VAAARGPLD---AATARSVALTISAYQEDPAYLRQCLASARALLYPPARLRVLMVDGN 137
Db 70 MRRAGR-PLKLCRSQRSRVALCIAAYQEDPEYLRKCLRSQAQRIAFP--NLKVVVMVDGN 126
QY 138 RAEDLYVMDVMEFVF-ADEDPATYVWDGNVHQPPEAAAGVAGAGAYREVEAEDPGRALV 196
Db 127 RQEDTYMLDIFHEVLGGTEQAGFFVWRNSFNFEAGEGETEASLQEGMR-----V 175
QY 197 EALVTRRCVCVAQRWGGKREVMYTAFAKALGSDVDYVQVCDSDTRLDPMALLELRVLD 256
Db 176 RAVVWASTFCIMQKMGKREVMYTAFAKALGSDVDYVQVCDSDTVLDPACTFEMLRVLE 235
QY 257 DPRVAVGGDVRIINPLDSWVSFLSSLRVYVAFNVERACQSYFHCVCISGPLESCPGPR 316
Db 236 DPQVGGVGGDVQILNKYDSWISFLSSVRYVWAFNVERACQSYFHCVCISGPLESCPGPR 316
QY 317 EHAMPSPFL 325
Db 292 RNSLLQOFL 300


```
RESULT 14
Q8CEB9          PRELIMINARY;          PRT;          554 AA.
ID Q8CEB9;
AC Q8CEB9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched
DE library, clone:4732404L04 product:similar to D942III.
GN Name=Has3;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=21085860; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690 (2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa Y., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kanakawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
```

```
RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK028582; BAC26017.1; -.
DR MGD; MGI:109599; Has3.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR InterPro; IPR001173; Glyco trans 2.
SQ SEQUENCE 554 AA; 63348 MW; BEFF239C901251DE CRC64;

Query Match 42.8%; Score 798; DB 2; Length 554;
Best Local Similarity 53.7%; Pred. No. 2.2e-58;
Matches 166; Conservative 40; Mismatches 81; Indels 22; Gaps 6;

Qy 21 RRVLTIFALLIIGLMTWYAAGVPLASDRYGLLAFLGFLYGAFLSAHLVAQSLFAYLEHRR 80
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
10 RVVGTSLFALVGLGILAAVYTGQFIHTEKHVLSFGLYGAILGLHLLIQSLFAFLHRR 69
Qy 81 VAAARAGPLD---AATARSVALTISAYQEDPAYLRQCLASARALLYPRARURVLWVDGN 137
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
70 MRRAGR-PLKHCQSQRPSVALCIAAYQEDPEYLRKCLRSQRIAFP--NLKVMVVDGN 126
Qy 138 RAEDLYMVDMPREVF-ADEDPATYVWDGNVHPWEPAAAGAVGAGAYREVAEDPGRUAV 196
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
127 RQEDTYMLDIFHEVLGTEQAGPFWRSPHEAGEGETEASLQEGMER-----V 175
Qy 197 EALVTRRCVCVAQRWGKGKREVMYTAFLGDSYDVYVQVCDSDTRLDPMALLELVRLDE 256
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
176 RAVVMASTFCIMQKQKREVMYTAFLGDSYDVYVQVCDSDTRLDPMALLELVRLDE 235
Qy 257 DPRVGAVGGDVRIINPLDSWVSLSLRYVAFVNERACQSYFHCVSCISGPLESCQPR 316
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
236 DPQVGCGVDVQLNKYDSWISFLSSVRYWAFVNERACQSYFHCVSCISGPLESCQPR 316
Qy 317 EHAMPSPFL 325
Db | | | | |
292 RNSLLQOFL 300

RESULT 15
Q8CH92          PRELIMINARY;          PRT;          554 AA.
ID Q8CH92;
AC Q8CH92;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hyaluronan synthase 3.
GN Name=HAS3;
OS Rattus norvegicus (Rat);
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14724275; DOI=10.1074/jbc.M313178200;
RA Itano N., Sawai T., Atsumi F., Miyashi O., Taniguchi S., Kannagi R.,
RA Hamaguchi M., Kimata K.;
RT "Selective expression and functional characteristics of three
RT mammalian hyaluronan synthases in oncogenic malignant
RT transformation.";
RL J. Biol. Chem. 279:18679-18687 (2004).
DR EMBL; AB097569; BAC43731.1; -.
DR InterPro; IPR001173; Glyco trans 2.
SQ SEQUENCE 554 AA; 63349 MW; 475DSBDG2717CE51 CRC64;

Query Match 42.8%; Score 798; DB 2; Length 554;
Best Local Similarity 53.7%; Pred. No. 2.2e-58;
Matches 166; Conservative 40; Mismatches 81; Indels 22; Gaps 6;

Qy 21 RRVLTIFALLIIGLMTWYAAGVPLASDRYGLLAFLGFLYGAFLSAHLVAQSLFAYLEHRR 80
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
10 RVVGTSLFALVGLGILAAVYTGQFIHTEKHVLSFGLYGAILGLHLLIQSLFAFLHRR 69
```

```

QY 81 VAAARGPLD---AATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRLMVVDGN 137
Db : | | | : | | | | | | | | | | | | | | | | | | | | | | | |
70 MRAGR-PLKLRCSQRRRSVALCIAAYQEDPEYLRKCLRSQRIAPP--NLKVMVVDGN 126
QY 138 RAEDLYMVDMPREVF-ADEDPATYVWDGNHQPWFEPAAAGAVGAGAYREVEAEDPRLAV 196
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
127 ROEDAYMLDIFHEVLGGTEQAGFFVWRSNFHEAGEGETEASLOEGMER-----V 175
QY 197 EALVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVOVCDSDTRLDPMALLELVRVLDE 256
Db : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
176 RAVWASTFSCIMQXWGGKREVMYTAFAKALGNSVDYIQVCDSDTVLDPACTIEMLRVLEE 235
QY 257 DPRVAVGGDVRIILNPLDSWVSFLSSLRYWVAFNVERACQSYFHCVSCISGPLESCGPR 316
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | |
236 DPQVGVGVDVQILNKYDSWISFLSSRYWMAFNVERACQSYFGCVQCISGFL---GMY 291
QY 317 EHAMPSEL 325
Db : | | | | |
292 RNSLIQQFL 300

```

Search completed: March 11, 2005, 14:21:58
 Job time : 96.0909 secs

THIS PAGE BLANK (US10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 14:11:23 ; Search time 26.25 Seconds
(without alignments)
1023.759 Million cell updates/sec

Title: US-10-672-399-6
Perfect score: 1864
Sequence: 1 MRQDAPKPTPAARRCSGLA.....ALVLPVAPADTLVQVLP 360

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents, AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1505	86.1	578	US-08-635-552A-2	Sequence 2, Appli
2	1597	85.7	589	US-09-949-016-11530	Sequence 11530, A
3	1512	81.1	583	US-08-675-499A-3	Sequence 3, Appli
4	1512	81.1	583	US-08-812-008-3	Sequence 3, Appli
5	1429	76.7	543	US-09-155-768-4	Sequence 4, Appli
6	804.5	43.2	587	US-08-635-552A-3	Sequence 3, Appli
7	800	42.9	554	US-08-812-008-32	Sequence 32, Appl
8	790.5	42.4	553	US-09-949-016-9599	Sequence 9599, Ap
9	788	42.3	587	US-08-675-499A-4	Sequence 4, Appli
10	788	42.3	587	US-08-812-008-4	Sequence 4, Appli
11	783.5	42.0	552	US-08-675-499A-2	Sequence 2, Appli
12	783.5	42.0	552	US-08-812-008-2	Sequence 2, Appli
13	783.5	42.0	552	US-09-949-016-6608	Sequence 6608, Ap
14	463.5	24.9	241	US-08-865-273-2	Sequence 2, Appli
15	463.5	24.9	241	US-09-385-174-2	Sequence 2, Appli
16	310	16.6	568	US-09-469-200E-10	Sequence 10, Appl
17	217	11.6	43	US-08-675-499A-9	Sequence 9, Appli
18	217	11.6	43	US-08-812-008-9	Sequence 9, Appli
19	204.5	11.0	426	US-08-675-499A-6	Sequence 6, Appli
20	204.5	11.0	426	US-08-812-008-6	Sequence 6, Appli
21	177	9.5	43	US-08-812-008-35	Sequence 35, Appli
22	174	9.3	417	US-09-469-200E-2	Sequence 2, Appli
23	173	9.3	43	US-08-675-499A-7	Sequence 7, Appli
24	173	9.3	43	US-08-812-008-7	Sequence 7, Appli
25	163	8.7	43	US-08-675-499A-10	Sequence 10, Appl
26	163	8.7	43	US-08-812-008-10	Sequence 10, Appl
27	151.5	8.1	419	US-08-270-581-2	Sequence 2, Appli

28	151.5	8.1	419	4	US-09-146-893-2	Sequence 2, Appli
29	151.5	8.1	419	4	US-08-675-499A-5	Sequence 5, Appli
30	151.5	8.1	419	4	US-08-812-008-5	Sequence 5, Appli
31	149.5	8.0	395	4	US-08-635-552A-4	Sequence 4, Appli
32	139	7.5	1093	4	US-09-248-796A-17108	Sequence 17108, A
33	135.5	7.3	393	4	US-09-902-540-11514	Sequence 11514, A
34	117.5	6.3	403	4	US-09-902-540-11529	Sequence 11529, A
35	110.5	5.9	904	4	US-09-252-991A-19257	Sequence 19257, A
36	104.5	5.6	419	4	US-09-328-352-7089	Sequence 7089, Ap
37	104.5	5.6	3519	3	US-09-428-517-4	Sequence 4, Appli
38	104	5.6	738	4	US-09-543-681A-7528	Sequence 7528, Ap
39	100.5	5.4	4150	3	US-09-428-517-2	Sequence 2, Appli
40	99.5	5.3	418	3	US-09-134-001C-4051	Sequence 4051, Ap
41	99.5	5.3	3562	4	US-09-679-279-14	Sequence 14, Appl
42	95.5	5.1	485	4	US-09-902-540-11019	Sequence 11019, A
43	95.5	5.1	1380	4	US-09-252-991A-25722	Sequence 25722, A
44	95	5.1	254	4	US-09-248-796A-16421	Sequence 16421, A
45	94	5.0	834	4	US-09-252-991A-24901	Sequence 24901, A

ALIGNMENTS

RESULT 1
US-08-635-552A-2
; Sequence 2, Application US/08635552A
; Patent No. 6423514
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; TITLE OF INVENTION: Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millicia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 578 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-635-552A-2

Query Match 86.1%; Score 1605; DB 4; Length 578;
Best Local Similarity 96.0%; Pred. No. 2.8e-179;
Matches 312; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

Qy 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLYLG 60
|||||
Db 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLYLG 60
|||||
Qy 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATFASVLTISAYQEDPAYLFQCLASARA 120
|||||

Db	61	AFLSAHLVAQSLFLAYLEHRRVAAAARGPLDPAATARSVALTTTSAYQEDPAYLRLQCLASAPA	120
Qy	121	LLYPRLRLVLMVMDGNGRAEDLYMVDMPREVFADDPATYYMDGNYHQHPWEPAAGAAGVA	180
Db	121	LLYPRLRLVLMVMDGNGRAEDLYMVDMPREVFADDPATYYMDGNYHQHPWEPAAGAAGVA	180
Qy	181	GAYREVEAEDPGLRLAVEALVTRTRCCVCAQRWGGKREVMYTAFKALGDSVDYVQVCDSDT	240
Db	181	GAYREVEAEDPGLRLAVEALVTRTRCCVCAQRWGGKREVMYTAFKALGDSVDYVQVCDSDT	240
Qy	241	RLDPMALLELVRLVLEDDPRVGAAGVDGVRILNPLDSWSVFLSSLRYWVAFNVERACQSIFYH	300
Db	241	RLDPMALLELVRLVLEDDPRVGAAGVDGVRILNPLDSWSVFLSSLRYWVAFNVERACQSIFYH	300
Qy	301	CVSCISGPLESCPGPREHAMMPSFL	325
Db	301	CVSCISGPL-----GLYRNNLLQQFL	321

RESULT 2
US-09-949-016-11530
; Sequence 11530, Application US/09949016
; Patent NO. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASES, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11530
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11530

```

RESULT 3
US-08-675-499A-3
; Sequence 3, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P. A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-675-499A-3

Query Match      81.1%; Score 1512; DB 4; Length 583
Best Local Similarity 90.6%; Pred. No. 2.4e-168;
Matches 298; Conservative 5; Mismatches 16; Indels

Qy 3 QDAPPTAARCSGLARRVLTIAPALLILGLMTWAYAAGVPLASDR
Db 2 RQDMPKPSAARCCSGLARRALTIIFALLILGLMTWAYAAGVPLASDR
Qy 63 LSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQV
Db 62 LSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQV
Qy 118 ARALLYPRARLRVLVMVDGNRAEDLYMVDMFREVFADEDPATYVWDGNT
Db 122 ARALLYPHTRLARVLVMVDGNRAEDLYMVDMFREVFADEDPATYVWDGNT
Qy 177 AVGAGAYRVEAEDPGLRAVEALVTRTRCVCAVQRWGKREVMYTAFAK
Db 182 AVGEGAYRVEAEDPGLRAVEALVTRTRCVCAVQRWGKREVMYTAFAK
Qy 237 DSDTRLDPMALLRLVRLDDEPRVGAAGVDVRLNPLDSWSVFLSSLR
Db 242 DSDTRLDPMALLRLVRLDDEPRVGAAGVDVRLNPLDSWSVFLSSLR
Qy 297 SYFHCVCISGPLESCPGPREHAMPSFL 325
Db 297 SYFHCVCISGPLESCPGPREHAMPSFL 325

```

Db 302 SYFHCVCISGSLP-----GLYRNNLLQQFL 326

RESULT 4

US-08-812-008-3
; Sequence 3, Application US/08812008
; Patent No. 6602693
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustine, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: U.S.A
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,008
; FILING DATE: 05-MAR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/675,499
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Embretson, Janet E
; REGISTRATION NUMBER: 39,665
; REFERENCE/DOCKET NUMBER: 150.183US1
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-812-008-3

Query Match 81.1%; Score 1512; DB 4; Length 583;
Best Local Similarity 90.6%; Pred. No. 2.4e-168;
Matches 298; Conservative 5; Mismatches 16; Indels 10; Gaps 3;

Qy 3 QODAPKPTPAARRCSGLARRVLTIAFALLIIGLMTWYAAGVPLASDRYGLLAFLYGF 62
Db 2 RQDMKPSAARCCSGLARRALTIFALLIIGLMTWYAAGVPLASDRYGLLAFLYGF 61
Qy 63 LSAHLVAQSLFAYLEHRRVAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCLAS 117
Db 62 LSAHLVAQSLFAYLEHRRVAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCLTS 121
Qy 118 ARALYPPARLRVLMVVDGNRAEDLYVMDPREVEFAEDPATYVWDGNVHQPWEPA-AAG 176
Db 122 ARALYPPHTRLRVLMVVDGNRAEDLYVMDPREVEFAEDPATYVWDGNVHQPWEPAATG 181
Qy 177 AVGAGAYREAEADPGRLAVALVTRRCVCVAQRWGKREVMYTAFAKALGSDVDYVQVC 236
Db 182 AVGAGAYREAEADPGRLAVALVTRRCVCVAQRWGKREVMYTAFAKALGSDVDYVQVC 241
Qy 237 DSDTRLDPMALLELRVRLDDEPRVAGVGDVRLNPLDSWSFSLSLRYWAFNVERACQ 296
Db 242 DSDTRLDPMALLELRVRLDDEPRVAGVGDVRLNPLDSWSFSLSLRYWAFNVERACQ 301

Qy 297 SYFHCVCISGPLESCPGPREHAMPSFL 325
Db 302 SYFHCVCISGSLP-----GLYRNNLLQQFL 326

RESULT 5

US-09-155-768-4
; Sequence 4, Application US/09155768A
; Patent No. 6162908
; GENERAL INFORMATION:
; APPLICANT: SEIKAGAKU CORPORATION
; TITLE OF INVENTION: POLYPEPTIDE OF HYALURONAN SYNTHASE OF
; FILE REFERENCE: HUMAN-ORIGIN AND DNA ENCODING THE SAME
; FILE REFERENCE: TOYAM37.001APC
; CURRENT APPLICATION NUMBER: US/09/155,768A
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: JP 8-084326
; EARLIER FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: JP 8-109663
; EARLIER FILING DATE: 1996-04-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 543
; TYPE: PRT
; ORGANISM: HUMAN
US-09-155-768-4

Query Match 76.7%; Score 1429; DB 3; Length 543;
Best Local Similarity 95.2%; Pred. No. 1.2e-158;
Matches 276; Conservative 4; Mismatches 6; Indels 4; Gaps 1;

Qy 36 MTWYAAGVPLASDRYGLLAFLYCAFLSAHLVAQSLFAYLEHRRVAAARGLDAATAR 95
Db 1 MTWYAAGVPLASDRYGLLAFLYCAFLSAHLVAQSLFAYLEHRRVAAARGLDAATAR 60
Qy 96 SVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAEDLYVMDPREVEFADE 155
Db 61 SVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAEDLYVMDPREVEFADE 120
Qy 156 DPATYVWDGNVHQPWEPAAGAGAYREAEADPGRLAVALVTRRCVCVAQRWGK 215
Db 121 DPATYVWDGNVHQPWEPAAGAGAYREAEADPGRLAVALVTRRCVCVAQRWGK 180
Qy 216 REVMTAFKALGSDVDYVQVCSDTRLDPMALLELRVRLDDEPRVAGVGDVRLNPLDS 275
Db 181 REVMTAFKALGSDVDYVQVCSDTRLDPMALLELRVRLDDEPRVAGVGDVRLNPLDS 240
Qy 276 WVSFLSLRYWAFNVERACQSYFHCVCISGPLESCPGPREHAMPSFL 325
Db 241 WVSFLSLRYWAFNVERACQSYFHCVCISGSLP-----GLYRNNLLQQFL 286

RESULT 6

US-08-635-552A-3
; Sequence 3, Application US/08635552A
; Patent No. 6423514
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; TITLE OF INVENTION: Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/635,552A
FILING DATE: 22-APR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LKS95-07
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 587 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-635-552A-3

Query Match 43.2%; Score 804.5; DB 4; Length 587;
Best Local Similarity 53.6%; Pred. No. 3.5e-85;
Matches 156; Conservative 51; Mismatches 75; Indels 9; Gaps 3;

QY 22 RVLTIAFALLILGLMTWAVAGVPLASDRYGLLAFGLYGAFLSAHLVAQSLFAYLEHRRV 81
DB RIYYSGVGLVLTATTAAYVAFQVLKHEALFSLGLYGLAWLLHMMQSLFAFLERVR 87
QY 82 AAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAED 141
DB 88 ---NKSELPCSKFKTVALTIAGYQENPEYLKLCLECKVVKPKDKXLIILVIDGNTDD 144
QY 142 LYMDMFREVFADEDPATYVMDGNHYQWPEPAAAGAVGAYREVEA---BDPGRLAVERA 198
DB 145 AYMMEFMKDFHGEDVGTGVWKNHYTVKKPE---ETNKGSCPEVSKPLNDEGINMVEE 201
QY 199 LVRTTRCVCVQWRGKREVMYTAFAKALGDSVDYVQVCDSDTRDPMALLELRVRLDDEP 258
DB 202 LVNRKRCVCIMQWQKREVMYTAFAQTGTSVDYVQVCDSDTKDELATVEMVKVLESND 261
QY 259 RVGAGGVDRILNPLDSDWVSFLSSLRVYVAFNVERACQSYFHCVSCISGGL 309
DB 262 MYGAGGVDRILNPNPDSIFSFSSLRVYVAFNVERACQSYFDCVSCISGGL 312

RESULT 7
US-08-812-008-32
Sequence 32, Application US/08812008
Patent No. 6602693
GENERAL INFORMATION:
APPLICANT: McDonald, J. A.
APPLICANT: Spicer, A. P.
APPLICANT: Augustine, M. L.
TITLE OF INVENTION: GENE ENCODING HYALURONAN
TITLE OF INVENTION: SYNTHASE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Moessner & Kluth, P.A.
STREET: P.O. Box 2938
CITY: Minneapolis
STATE: MN
COUNTRY: U.S.A.
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/812,008
FILING DATE: 05-MAR-1997

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/675,499
FILING DATE: 03-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Embretson, Janet E.
REGISTRATION NUMBER: 39,665
REFERENCE/DOCKET NUMBER: 150.183US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-373-6900
TELEFAX: 612-339-3061
TELEX:
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 554 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-812-008-32

Query Match 42.9%; Score 800; DB 4; Length 554;
Best Local Similarity 53.7%; Pred. No. 1.1e-84;
Matches 166; Conservative 41; Mismatches 80; Indels 22; Gaps 6;

QY 21 RRVLTIAFALLILGLMTWAVAGVPLASDRYGLLAFGLYGAFLSAHLVAQSLFAYLEHRR 80
DB 10 RVGTSLFALVVLGGILAAVYTGQFIHTEKHYSFGLYGAILGHLHLLIQSLFAPLEHRR 69
QY 81 VAAARGPLD---AATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGN 137
DB 70 MRAGR-PLKLHCQSRSRVALCIAAQEDPEYLKCLRSAQRIAPP--NLKVVVVVDGN 126
QY 138 RAEDLYMVDMFREVF-ADDPATYVMDGNHYQWPEPAAAGAVGAYREVEAEDPGRVAV 196
DB 127 RQEDTYMLDIFHEVLGGTEQAQGFVWRSNFHEAGEGETEASLQEGMER-----V 175
QY 197 EALVTRCVCVQWRGKREVMYTAFAKALGDSVDYVQVCDSDTRDPMALLELRVRLDDE 256
DB 176 RAVVWASTFTCMQWQKREVMYTAFAKALGNSVDYIQVCDSDTVLPDPACTIEMLRVLEE 235
QY 257 DPRVAGGVDRILNPLDSDWVSFLSSLRVYVAFNVERACQSYFHCVSCISGFLSCPGPR 316
DB 236 DPQVGGVGGDVQILNKYDSNIFLSSRVYVWAFNVERACQSYFCVCQICISGPL----GMY 291
QY 317 EHMMPSFL 325
DB 292 RNSLLQQFL 300

RESULT 8
US-09-949-016-9599
Sequence 9599, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9599
LENGTH: 553
TYPE: PRT


```

; ORGANISM: Human
US-09-949-016-9599

Query Match          42.4%; Score 790.5; DB 4; Length 553;
Best Local Similarity 53.4%; Pred No. 1.4e-83;
Matches 164; Conservative 39; Mismatches 85; Indels 19; Gaps 5;

QY 21 RRVLTAFALLILGLMTWAYAAGVPLASDRYGLLAGFLYGAPLSAHLVAQSLFAYLEHRR 80
DB 10 RVVGTSLFALAVLGILAAVTVGYQFIHTEKHYSFLGYLAILGLHLLIQLSLFAPLEHRR 69

QY 81 VAAARAG-PLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRA 139
DB 70 MRRAGQALKLSPRRGSVALCIAAYQEDPDYLRKCLRSQAQRISFP--DLKVMVMVVDGNRQ 127

QY 140 EDLYVMDVMEFVFADEDPATVWDGNVHQPWEPAAGAGVAGAVEAEDPGRLAVEA 198
DB 128 EDAYMLDIFHEVLGTEQANGFVWRNSFHEAGEGETEASLQEGMDR-----VRD 176

QY 199 LVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDTRLDPMALLLVRVLDSDP 258
DB 177 VVRASTFSCIMQKWKGGKREVMYTAFAKALGDSVDYVQVCDSDTVLDPACTIEMLRVLESDP 236

QY 259 RVGAVGGDVRILNPLDSWVSFLSLRYWVAFNVERACOSYFHCVSCISGPGPREH 318
DB 237 QVGGVGGDVQILNKYDSWISFLSVRYWMAFNVERACOSYFGCVQICISGPL-----GMVYN 292

QY 319 AMPSFL 325
DB 293 SLQQFL 299
```

```

RESULT 9
US-08-675-499A-4
; Sequence 4, Application US/08675499A
; Patent No. 6452150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
```

```

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-675-499A-4

Query Match          42.3%; Score 788; DB 4; Length 587;
Best Local Similarity 53.3%; Pred. No. 3.1e-83;
Matches 155; Conservative 51; Mismatches 75; Indels 10; Gaps 4;

QY 22 RVLTIAPALLILGLMTWAYAAGVPLASDRYGLLAGFLYGAPLSAHLVAQSLFAYLEHRRV 81
DB 28 RIITYSGVLLATITAAVVAEFOVLKHEATLFLSLGLYGLAMLLHMMQSLFAFLEIRRV 87

QY 82 AAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAED 141
DB 88 ---NKSELPSCFKTKTVALTIAGYQENPEYLLIKLESCKYVKYPKDKLILVIDGNTEDD 144

QY 142 LYVDMFREVFADEDPATVWDGNVHQPWEPAAGAGVAGAVEA---EDPGRLAVEA 198
DB 145 AYMEMFKDVPFHGSDVGTVMKGNVHTVKKPE---ETNKGSCPEVSKPLNEDEGINMVEE 201

QY 199 LVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDTRLDPMALLLVRVLDSDP 258
DB 202 LVRNKRVCVIMQQW--GKREVMYTAFAIGTSVDYVQVCDSDTKLDELATVEMKVLESND 260

QY 259 RVGAVGGDVRILNPLDSWVSFLSLRYWVAFNVERACOSYFHCVSCISGSL 309
DB 261 MYGAVGGDVRILNPLDYSFISFNSRLRYWMAFNVERACOSYFDCVSCISGSL 311
```

```

RESULT 10
US-08-812-008-4
; Sequence 4, Application US/08812008
; Patent No. 6602693
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: U.S.A
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,008
; FILING DATE: 05-MAR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/675,499
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Embretson, Janet E
; REGISTRATION NUMBER: 39,665
; REFERENCE/DOCKET NUMBER: 150.183US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
```

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-812-008-4

Query Match      42.3%; Score 788; DB 4; Length 587;
Best Local Similarity 53.3%; Pred. No. 3.1e-83;
Matches 155; Conservative 51; Mismatches 75; Indels 10; Gaps 4;

Qy 22 RVLTIAPALLILGLMTWYAAAGVP-LASDRYGLLAGLGLYGAFLSAHLVAQSLFAYLEHRRV 81
Db 28 RIIVYFSGVLLATITAAAYVAEFQVLKHEALIFSLGLYLGLAMLLHMMQSLFAFLFLEIRRV 87
Qy 82 AAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRLRLVLMVVDGNRAED 141
Db 88 ---NKSELPFCFKTVALTIAGYQENPEYLTKLESCKYVYKPKGLKIIILVIDGNTEDD 144
Qy 142 LYMDMFREVADEDPATYVWDGNHQPWEPAAGAVGAGAYREVEA---EDPGRLAVEA 198
Db 145 AYMEMFQDVFGEDVGTVMKGNHYTVKKEP---ETNKGSCPEVSKPLNEDEGINNVVEE 201
Qy 199 LVTRRCVCVAQRMGKREVMYTAFAKALGDSVYVQVCDSDTRLDPMALLELRVLDSDP 258
Db 202 LVNRKRCVICMQW-GKREVMYTAFAIGTSVDYVQVCDSDTKLDELATVEMVVKYLESD 260
Qy 259 RVGAVGGDVRLNPLDSWVSLSSLRVWAFNVERACOSYFHCVSCISGGL 309
Db 261 MYGAVGGDVRLNPNYDPSIFSMSSLRVWMAFNVERACOSYFDCVSCISGGL 311

RESULT 11
US-08-675-499A-2
; Sequence 2, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-675-499A-2

Query Match      42.0%; Score 783.5; DB 4; Length 552;
Best Local Similarity 51.3%; Pred. No. 9.4e-83;
Matches 157; Conservative 52; Mismatches 76; Indels 21; Gaps 7;

Qy 21 RRVLTIAFALLILGLMTWYAAAGVP-LASDRYGLLAGLGLYGAFLSAHLVAQSLFAYLEHR 79
Db 11 RIICITTLFGVLLGITAAAYIVGQFOTDNY-YFSFGLYGAFLASHLIIOSLFAFLFLEHR 69
Qy 80 RVAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRLRLVLMVVDGNRA 139
Db 70 KMKSLSTFPI--KLNTKVALCIAAYQEDPDYLRKCLQSVKELTYP--GIKVMVWIDGNSD 125
Qy 140 EDLYMDMFREVADEDPATYVWDGNHQPWEPAAGAVGAGAYREVEAEDPGRLAVEAL 199
Db 126 DDLYMDIFSEVIGRDKSATYVKNPFHEK-----GPGETEESHKSSQH--VTQL 174
Qy 200 VTRRCVCVAQRMGKREVMYTAFAKALGDSVYVQVCDSDTRLDPMALLELRVLDSDP 259
Db 175 VLSNKSICIMQKMGKREVMYTAFAKALGDSVYVQVCDSDTMDLPASSVEMVKVLEBDPM 234
Qy 260 VGAVGGDVRLNPLDSWVSLSSLRVWAFNVERACOSYFHCVSCISGPLESCPCPREHA 319
Db 235 VGVGGDVQVILNKYDSWISFLSSRVWMAFNVERACOSYFGCVQICISGPL-----GMYRNS 290
Qy 320 MPPSFL 325
Db 291 LLHEFV 296

RESULT 12
US-08-812-008-2
; Sequence 2, Application US/08812008
; Patent No. 6602693
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: U.S.A
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,008
; FILING DATE: 05-MAR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/675,499
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Embretson, Janet E
; REGISTRATION NUMBER: 39,665
; REFERENCE/DOCKET NUMBER: 150.183US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
```

Qy	138	RAEDLYWDMFREVFADDPATYVNDGNVQWPEAPAAAGAVGAGAYREVEAEDPRLAVE	19
Db	124	SEDDLYMMDIFSEVWGRDKSATYIWKNFHEK-----GPGETDESHKSSQH--VT	172
Qy	198	ALVTRRCVCVAORWGGKREVMYTAFAKALGDSVDYVQVCDSDTRLDPMALLEVRLVLD	257
Db	173	QVLSNKSICIMQKNGKREVMYTAFAKALGDSVDYVQVCDSDTRLDPMALLEVRLVLD	232
Qy	258	PRVGVGGDVRLNPLDSDWVSLSSLYRYVAFNVERACOSYFHCVSCISGPLESCPGPRE	317
Db	233	PMVGGVGGDVQLLNKYDSWISFLSSVRYWMAFNIERACOSYFGCVQCISGPLE---	288
Qy	318	HAMPSFL 325	
Db	289	NSLHEFV 296	
<p>RESULT 14</p> <p>US-08-865-273-2</p> <p>Sequence 2, Application US/08865273</p> <p>Patent No. 5994100</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: ZHU, YUAN</p> <p>APPLICANT: NAMEI, PORNAL</p> <p>APPLICANT: PULLEN, MARK A</p> <p>TITLE OF INVENTION: NOVEL HAS2 SPLICING VARIANT</p> <p>TITLE OF INVENTION: NOVEL HAS2 SPLICING VARIANT</p> <p>NUMBER OF SEQUENCES: 3</p> <p>CORRESPONDENCE ADDRESS:</p> <p>ADDRESSEE: RATHER & PRESTIA</p> <p>STREET: P.O. BOX 980</p> <p>CITY: VALLEY FORGE</p> <p>STATE: PA</p> <p>COUNTRY: USA</p> <p>ZIP: 19482</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Diskette</p> <p>COMPUTER: IBM Compatible</p> <p>OPERATING SYSTEM: DOS</p> <p>SOFTWARE: FastSEQ for Windows Version 2.0</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/865,273</p> <p>FILING DATE: 29-MAY-1997</p> <p>CLASSIFICATION: 435</p> <p>PRIOR APPLICATION DATA:</p> <p>APPLICATION NUMBER:</p> <p>FILING DATE:</p> <p>ATTORNEY/AGENT INFORMATION:</p> <p>NAME: PRESTIA, PAUL F</p> <p>REGISTRATION NUMBER: 23,031</p> <p>REFERENCE/DOCKET NUMBER: GH-70053</p> <p>TELECOMMUNICATION INFORMATION:</p> <p>TELEPHONE: 610-407-0700</p> <p>TELEFAX: 610-407-0701</p> <p>TELEX: 846169</p> <p>INFORMATION FOR SEQ ID NO: 2:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 241 amino acids</p> <p>TYPE: amino acid</p> <p>STRANDEDNESS: single</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: protein</p> <p>US-08-865-273-2</p>			
<p>Query Match 24.9%; Score 463.5; DB 2; Length 241;</p> <p>Best Local Similarity 45.4%; Pred. No. 1e-45;</p> <p>Matches 99; Conservative 39; Mismatches 63; Indels 17; Gaps 6</p>			
Qy	19	LARRVLTIAFALLILGLMTWYAAGVP--LASDRYGLLAFGLYGAFLSAHLVAQSLFAYLE	77
Db	9	ILRIIGTTLFGVSLGITAAYVGVQFIQTDNY--YFSEGLYGAFLASHLIQSLEFALE	67
Qy	78	HRVAAARGPLDPAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGN	137
Db	68	HRKWKSLTETPI--KLNTKTVLCIAAYQEDPDYLRKCLQSVKRLTYP--GIKVVNVVDGN	123

Db 68 HRKMKSLPTI--KLNKTVLALCIAAYQEDPDYLRKCLQSVKRLTYP--GIKVVNVIDGN 123
QY 138 RAEDLYMVDMEVFADDPATYVWDGNYHQPWEPAAGAGAGAYREVEAEDPGRLAVE 197
Db 124 SEDDLYMMDIFSEVWGRDKSATHIWKNNFHEK-----GPGETDESHKSSQH--VT 172
QY 198 ALVTRTRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQV 235
Db 173 QLVLSNKSICIMQKWKGGKREVMYTAFAKALGDSVDYVQV 210

RESULT 15

US-09-385-174-2
; Sequence 2, Application US/09385174
; Patent No. 6350446
; GENERAL INFORMATION:
; APPLICANT: ZHU, YUAN
; NAMBI, PONNAL
; PULLEN, MARK A
; TITLE OF INVENTION: NOVEL HAS2 SPLICING VARIANT
; HOEFC11: A TARGET IN CHRONIC RENAL FAILURE,
; INFLAMMATORY DISEASES AND MYOCARDIAL ISCHEMIA
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/385,174
; FILING DATE: 30-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/865,273
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 241 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-385-174-2

Query Match 24.9%; Score 463.5; DB 3; Length 241;
Best Local Similarity 45.4%; Pred. No. 1e-45;
Matches 99; Conservative 39; Mismatches 63; Indels 17; Gaps 6;
QY 19 LARRVLTIAFALLIIGLMTWAYAGVP-LASDRYGLLAFLGYGAPLSAHLVAQSLFAYLE 77
Db 9 ILRIIGTTLFGVSLGLLGTAAIVGVQFIQTDNY-YFSFGLYGAFSLASHLIQSLFAYLE 67
QY 78 HRRVAAAARGPDAAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVDGN 137
Db 68 HRKMKSLPTI--KLNKTVLALCIAAYQEDPDYLRKCLQSVKRLTYP--GIKVVNVIDGN 123

QY 138 RAEDLYMVDMEVFADDPATYVWDGNYHQPWEPAAGAGAGAYREVEAEDPGRLAVE 197
Db 124 SEDDLYMMDIFSEVWGRDKSATHIWKNNFHEK-----GPGETDESHKSSQH--VT 172
QY 198 ALVTRTRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQV 235
Db 173 QLVLSNKSICIMQKWKGGKREVMYTAFAKALGDSVDYVQV 210

Search completed: March 11, 2005, 14:24:37
Job time : 31.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 14:22:15 ; Search time 80.7955 Seconds
(without alignments)
1469.688 Million cell updates/sec

Title: US-10-672-399-6
Perfect score: 1864
Sequence: 1 MRQDAPKPTPAARRCSGLA.....ALVPAVAEPADTLVQVVLP 360

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1396920 seqs, 329844858 residues

Total number of hits satisfying chosen parameters: 1396920

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1864	100.0	360	16	US-10-672-399-6
2	1612	86.5	376	16	US-10-672-399-4
3	1605	86.1	578	13	US-10-042-523-2
4	1605	86.1	578	16	US-10-672-399-2
5	1512	81.1	583	10	US-09-902-939-4
6	1298	69.6	320	16	US-10-672-399-8
7	804.5	43.2	587	13	US-10-042-523-3
8	804.5	43.2	588	15	US-10-309-560-10
9	800	42.9	554	10	US-09-902-939-5
10	790.5	42.4	553	15	US-10-295-027-370
11	790.5	42.4	553	15	US-10-188-832-137
12	783.5	42.0	552	10	US-09-902-939-5
13	783.5	42.0	552	14	US-10-262-526-2

14	783.5	42.0	552	14	US-10-262-526-4	Sequence 4, Appli
15	310	16.6	567	9	US-09-879-959-7	Sequence 7, Appli
16	310	16.6	567	14	US-10-172-527-7	Sequence 7, Appli
17	310	16.6	567	15	US-10-309-560-12	Sequence 12, Appli
18	310	16.6	568	14	US-10-011-768B-10	Sequence 10, Appli
19	310	16.6	568	14	US-10-011-771B-10	Sequence 10, Appli
20	174	9.3	417	9	US-09-879-959-2	Sequence 2, Appli
21	174	9.3	417	14	US-10-011-768B-2	Sequence 2, Appli
22	174	9.3	417	14	US-10-011-771B-2	Sequence 2, Appli
23	174	9.3	417	14	US-10-172-527-2	Sequence 2, Appli
24	174	9.3	417	14	US-10-326-185-2	Sequence 2, Appli
25	174	9.3	417	15	US-10-309-560-2	Sequence 2, Appli
26	174	9.3	417	15	US-10-309-560-16	Sequence 16, Appli
27	174	9.3	417	15	US-10-309-560-17	Sequence 17, Appli
28	174	9.3	417	15	US-10-309-560-18	Sequence 18, Appli
29	174	9.3	417	15	US-10-309-560-22	Sequence 22, Appli
30	174	9.3	417	15	US-10-309-560-23	Sequence 23, Appli
31	174	9.3	417	15	US-10-309-560-24	Sequence 24, Appli
32	174	9.3	417	15	US-10-309-560-25	Sequence 25, Appli
33	170	9.1	417	15	US-10-309-560-15	Sequence 15, Appli
34	170	9.1	417	15	US-10-309-560-19	Sequence 19, Appli
35	170	9.1	417	15	US-10-309-560-20	Sequence 20, Appli
36	170	9.1	417	15	US-10-309-560-21	Sequence 21, Appli
37	170	9.1	417	15	US-10-309-560-26	Sequence 26, Appli
38	170	9.1	417	15	US-10-309-560-27	Sequence 27, Appli
39	170	9.1	417	15	US-10-309-560-28	Sequence 28, Appli
40	170	9.1	417	15	US-10-309-560-29	Sequence 29, Appli
41	159	8.5	1219	15	US-10-389-493-3212	Sequence 3212, Ap
42	156.5	8.4	1165	9	US-09-801-368-76	Sequence 76, Appli
43	156.5	8.4	1165	15	US-10-369-493-1432	Sequence 1432, Ap
44	156	8.4	1160	15	US-10-369-493-12860	Sequence 12860, A
45	152.5	8.2	419	15	US-10-309-560-30	Sequence 30, Appli

ALIGNMENTS

RESULT 1
US-10-672-399-6
; Sequence 6, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-6

Query Match 100.0%; Score 1864; DB 16; Length 360;
Best Local Similarity 100.0%; Pred. No. 2,8e-178;
Matches 360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAGVPLASDRYGLLAFLYLG	60
Db	1	MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAGVPLASDRYGLLAFLYLG	60
Qy	61	AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAAATARSVALTISAYQEDPAYLRQCLASARA	120
Db	61	AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAAATARSVALTISAYQEDPAYLRQCLASARA	120
Qy	121	LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQPWEPAAAGAVGA	180
Db	121	LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQPWEPAAAGAVGA	180

```

QY 181 GAYREVEAEDPGLAVEALVRTRRCVCAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db 181 GAYREVEAEDPGLAVEALVRTRRCVCAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
QY 241 RLDPMALLELRVRLDDEPRVGAAGVGVRLNPLDLSWVSLSLRYWVAFNVERACQSYFH 300
Db 241 RLDPMALLELRVRLDDEPRVGAAGVGVRLNPLDLSWVSLSLRYWVAFNVERACQSYFH 300
QY 301 CVSCISGPLESCPGPREHAMPSFLAPVQVHLQVPLLRLDALVLPVAVAPADTLVQVVL 360
Db 301 CVSCISGPLESCPGPREHAMPSFLAPVQVHLQVPLLRLDALVLPVAVAPADTLVQVVL 360

```

RESULT 2

```

US-10-672-399-4
; Sequence 4, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 4
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-4

```

```

Query Match      86.5%; Score 1612; DB 16; Length 376;
Best Local Similarity 98.7%; Pred. No. 6e-153;
Matches 311; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

QY 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYL 60
Db 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYL 60
QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILASARA 120
Db 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILASARA 120
QY 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQWPPEAAGAVGA 180
Db 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQWPPEAAGAVGA 180
QY 181 GAYREVEAEDPGLAVEALVRTRRCVCAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db 181 GAYREVEAEDPGLAVEALVRTRRCVCAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
QY 241 RLDPMALLELRVRLDDEPRVGAAGVGVRLNPLDLSWVSLSLRYWVAFNVERACQSYFH 300
Db 241 RLDPMALLELRVRLDDEPRVGAAGVGVRLNPLDLSWVSLSLRYWVAFNVERACQSYFH 300
QY 301 CVSCISGPLESCPGP 315
Db 301 CVSCISGPLESCPGP 315

```

RESULT 3

```

US-10-042-523-2
; Sequence 2, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

```

```

; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/042,523
; FILING DATE: 19-Oct-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 578 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-042-523-2

```

```

Query Match      86.1%; Score 1605; DB 13; Length 578;
Best Local Similarity 96.0%; Pred. No. 5.5e-152;
Matches 312; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

```

```

QY 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYL 60
Db 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWAYAAGVPLASDRYGLLAFLGYL 60
QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILASARA 120
Db 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCILASARA 120
QY 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQWPPEAAGAVGA 180
Db 121 LLYPRARLRVLMVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQWPPEAAGAVGA 180
QY 181 GAYREVEAEDPGLAVEALVRTRRCVCAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Db 181 GAYREVEAEDPGLAVEALVRTRRCVCAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
QY 241 RLDPMALLELRVRLDDEPRVGAAGVGVRLNPLDLSWVSLSLRYWVAFNVERACQSYFH 300
Db 241 RLDPMALLELRVRLDDEPRVGAAGVGVRLNPLDLSWVSLSLRYWVAFNVERACQSYFH 300
QY 301 CVSCISGPLESCPGPREHAMPSFL 325
Db 301 CVSCISGPLESCPGPREHAMPSFL 325

```

RESULT 4

```

US-10-672-399-2
; Sequence 2, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401

```

```
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-2

Query Match      86.1%; Score 1605; DB 16; Length 578;
Best Local Similarity 96.0%; Pred. No. 5 Se-152;
Matches 312; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

QY 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFGLYG 60
DB 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFGLYG 60
QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
DB 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
QY 121 LLYPRARLRVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAGAVGA 180
DB 121 LLYPRARLRVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAGAVGA 180
QY 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMYTAFAKALGSDVDYVQVCDSDT 240
DB 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMYTAFAKALGSDVDYVQVCDSDT 240
QY 241 RLDPMALLELVRLVDEDPGVAGGDDVRLINPLDSWVSFLSLRYWVAFNVERACQSYFH 300
DB 241 RLDPMALLELVRLVDEDPGVAGGDDVRLINPLDSWVSFLSLRYWVAFNVERACQSYFH 300
QY 301 CVSCISGPLESCPGPREHAMMPSFL 325
DB 301 CVSCISGPLESCPGPREHAMMPSFL 325
US-10-672-399-3

Query Match      81.1%; Score 1512; DB 10; Length 583;
Best Local Similarity 90.6%; Pred. No. 1.2e-142;
Matches 298; Conservative 5; Mismatches 16; Indels 10; Gaps 3;

QY 3 QODAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFGLYGAF 62
DB 2 QODMPKPEARCCSGSLARRALTIFALLILGLMTWYAAGVPLASDRYGLLAFGLYGAF 61
QY 63 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLAS 117
DB 62 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLTS 121
QY 118 ARALLYPRARLRVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAG 176
DB 122 ARALLYPHTRUKVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAG 181

; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-2

Query Match      86.1%; Score 1605; DB 16; Length 578;
Best Local Similarity 96.0%; Pred. No. 5 Se-152;
Matches 312; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

QY 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFGLYG 60
DB 1 MRQDAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFGLYG 60
QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
DB 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
QY 121 LLYPRARLRVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAGAVGA 180
DB 121 LLYPRARLRVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAGAVGA 180
QY 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMYTAFAKALGSDVDYVQVCDSDT 240
DB 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMYTAFAKALGSDVDYVQVCDSDT 240
QY 241 RLDPMALLELVRLVDEDPGVAGGDDVRLINPLDSWVSFLSLRYWVAFNVERACQSYFH 300
DB 241 RLDPMALLELVRLVDEDPGVAGGDDVRLINPLDSWVSFLSLRYWVAFNVERACQSYFH 300
QY 301 CVSCISGPLESCPGPREHAMMPSFL 325
DB 301 CVSCISGPLESCPGPREHAMMPSFL 325
US-10-672-399-3

Query Match      81.1%; Score 1512; DB 10; Length 583;
Best Local Similarity 90.6%; Pred. No. 1.2e-142;
Matches 298; Conservative 5; Mismatches 16; Indels 10; Gaps 3;

QY 3 QODAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAGVPLASDRYGLLAFGLYGAF 62
DB 2 QODMPKPEARCCSGSLARRALTIFALLILGLMTWYAAGVPLASDRYGLLAFGLYGAF 61
QY 63 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLAS 117
DB 62 LSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLTS 121
QY 118 ARALLYPRARLRVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAG 176
DB 122 ARALLYPHTRUKVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAG 181

; Sequence 4, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHazy
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/04020-USO
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 583
; TYPE: PRT
; ORGANISM: mouse
US-09-902-939-4

; Sequence 4, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 8
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-8

Query Match      69.6%; Score 1298; DB 16; Length 320;
Best Local Similarity 95.1%; Pred. No. 1.6e-121;
Matches 252; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

QY 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
DB 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
QY 121 LLYPRARLRVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAGAVGA 180
DB 61 LLYPRARLRVLMVVDGNRAEDLYMVDMPREVEFADEDPATYVWDGNHQPWEPAAAGAVGA 120
QY 181 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMYTAFAKALGSDVDYVQVCDSDT 240
DB 121 GAYREVEAEDPGLAVEALVTRRCVCVQAQRWGGKREVMYTAFAKALGSDVDYVQVCDSDT 180
QY 241 RLDPMALLELVRLVDEDPGVAGGDDVRLINPLDSWVSFLSLRYWVAFNVERACQSYFH 300
DB 181 RLDPMALLELVRLVDEDPGVAGGDDVRLINPLDSWVSFLSLRYWVAFNVERACQSYFH 240
QY 301 CVSCISGPLESCPGPREHAMMPSFL 325
DB 241 CVSCISGPLESCPGPREHAMMPSFL 261

; Sequence 3, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; CORRESPONDENCE ADDRESS:
; NUMBER OF SEQUENCES: 4
; ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
```



```
Db          292 RNSLQQFL 300
RESULT 10
US-10-295-027-370
; Sequence 370, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gieh, Kurt C.
; APPLICANT: Glyme, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 370
; TYPE: PRT
; LENGTH: 553
; ORGANISM: Homo sapiens
US-10-295-027-370

Query Match          42.4%; Score 790.5; DB 15; Length 553;
Best Local Similarity 53.4%; Pred. No. 3.2e-70;
Matches 164; Conservative 39; Mismatches 85; Indels 19; Gaps 5;

QY      21  RRVLTAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYGAFLSAHLVAQSLPAYLEHRR 80
DB      10  RVVGTSLFALAVLGGILAAVYTGQFIHTEKHYSFGLYGAILGHLILIQSLFALEHRR 69

QY      81  VAAAARG-PLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRABLRVLMVVDGNRA 139
DB      70  MRRAGQALKPSPRRGSVALCIAAYQEDPDYLRKCLRSQAQRISFP--DLKVMVMVVDGNRQ 127

QY      140 EDLYVMDMFREVF-ADEDPATYVWDGNYHQWPEPAAGAVGAGAYREVEADPGRLAYEA 198
DB      128 EDAYMLDIFHEVLGGTEQAGFFVMRSNFHEAGEGETEASLQEGMDR-----VRD 176

QY      199 LVRTTRRCVCAQRWGKGKREVMYTAFAKLGDSVDYVQVCDSDTRLDPMALLELRVLDSDP 258
DB      177 VVRSTFSCIQKQKGGKREVMYTAFAKLGDSVDYIQVCDSDTVLPDPACTIEMLRVLEEDP 236

QY      259 RVGAVGGDVRILNPLDSWVSFLSSLRVYVAFNVERACOSYFHCVSCISGPLESCEPGPREH 318
DB      237 QVGGVGGDVQILNKYDSWISFLSSVRVYVWMAFNVERACOSYFGCVQCISGPL-----GMVRN 292

QY      319 AMMPSFL 325
DB      293 SLLOQFL 299

RESULT 12
US-09-902-939-5
; Sequence 5, Application US/09902939
; Publication No. US20030087850A1

Db          292 RNSLQQFL 300
RESULT 11
US-10-188-832-137
; Sequence 137, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Nataasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 137
; TYPE: PRT
; LENGTH: 553
; ORGANISM: Homo sapiens
US-10-188-832-137

Query Match          42.4%; Score 790.5; DB 15; Length 553;
Best Local Similarity 53.4%; Pred. No. 3.2e-70;
Matches 164; Conservative 39; Mismatches 85; Indels 19; Gaps 5;

QY      21  RRVLTAFALLILGLMTWAYAAGVPLASDRYGLLAFGLYGAFLSAHLVAQSLPAYLEHRR 80
DB      10  RVVGTSLFALAVLGGILAAVYTGQFIHTEKHYSFGLYGAILGHLILIQSLFALEHRR 69

QY      81  VAAAARG-PLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRABLRVLMVVDGNRA 139
DB      70  MRRAGQALKPSPRRGSVALCIAAYQEDPDYLRKCLRSQAQRISFP--DLKVMVMVVDGNRQ 127

QY      140 EDLYVMDMFREVF-ADEDPATYVWDGNYHQWPEPAAGAVGAGAYREVEADPGRLAYEA 198
DB      128 EDAYMLDIFHEVLGGTEQAGFFVMRSNFHEAGEGETEASLQEGMDR-----VRD 176

QY      199 LVRTTRRCVCAQRWGKGKREVMYTAFAKLGDSVDYVQVCDSDTRLDPMALLELRVLDSDP 258
DB      177 VVRSTFSCIQKQKGGKREVMYTAFAKLGDSVDYIQVCDSDTVLPDPACTIEMLRVLEEDP 236

QY      259 RVGAVGGDVRILNPLDSWVSFLSSLRVYVAFNVERACOSYFHCVSCISGPLESCEPGPREH 318
DB      237 QVGGVGGDVQILNKYDSWISFLSSVRVYVWMAFNVERACOSYFGCVQCISGPL-----GMVRN 292

QY      319 AMMPSFL 325
DB      293 SLLOQFL 299
```



```
; Sequence 7, Application US/09879959
; Patent No. US20020160489A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H
; APPLICANT: Kumari, Kshama
; APPLICANT: DeAngelis, Paul
; TITLE OF INVENTION: STREPTOCOCCUS EQUISMILIS HYALURONAN SYNTHASE GENE AND EXPRESSION
; FILE REFERENCE: 3554.049
; CURRENT APPLICATION NUMBER: US/09/879,959
; CURRENT FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: 09/469,200
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: 09/178,851
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 567
; TYPE: PRT
; ORGANISM: Chlorella virus PBCV-1
US-09-879-959-7
```

```
Query Match      16.6%; Score 310; DB 9; Length 567;
Best Local Similarity 29.4%; Pred. No. 5.8e-22;
Matches 89; Conservative 57; Mismatches 107; Indels 50; Gaps 10;

QY 19 LARRVLTITAFALLIIG-----LMTWAVAAGVPLASDRYGLLAFGLYGAFTLSAHLVAQSL 72
DB 15 ITSNIAGGASLLIAPAITGYVLHWNIALSTINGVSAYGIFVGF-----LAQVL 66
QY 73 FAYLEHRR-----VAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRRL 128
DB 67 FSELNRKRLRWISLRPKGNDV----RLAVIAGVREDPYWFKCLESVRDSYGNV-A 121
QY 129 RVLMTVDGNRAEDLYWDMFREVFADEDPATYWDGNTHQWPAPAAAGAGAYREVEA 188
DB 122 RLICVIDGDEDDMRMAAVYKAIYND-----NIKPP-----EFVLCE 159
QY 189 EDPGRLAVALVTRRCVCVQORWGKREVMYTFK--ALGDSVDYVQVCDSDTRLDPMA 246
DB 160 DKEGERIDS--DFSRDICVLQPHRGKRECLTGFQLAKMDPSVNAVLIIDSDTVLEKDA 217
QY 247 LLELYRVLDEDPVGAAGDVRIINPLDSWSVFLSRLRYWAFNVERACQSYFHCVCSCIS 306
DB 218 ILEVYVPLACDPEIQAVAGECKIWN-TDILLSLLVAVRYSAFCVERSAQSFRTVQCVG 276
QY 307 GPL 309
DB 277 GPL 279
```

Search completed: March 11, 2005, 14:44:38
Job time : 81.7955 secs

THIS PAGE BLANK (SP10)

Db	636	AGGACTCGCAGTGTGCTGCTGGCGCAGCGCTGGGGCGGCAAGCGGAGTCTATGTAC	695
Qy	661	ACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACACA	720
Db	696	ACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACACA	755
Qy	721	AGGTTGGACCCCATGCGACTGCTGGAGCTGTGCGGGTACTGGACGAGACCCCGGGTA	780
Db	756	AGGTTGGACCCCATGCGACTGCTGGAGCTGTGCGGGTACTGGACGAGACCCCGGGTA	815
Qy	781	GGGGCTGTGGTGGGAGCTGGGATCCCTTAACCTCTGGACTCCCTGGGTGAGCTTCCTA	840
Db	816	GGGGCTGTGGTGGGAGCTGGGATCCCTTAACCTCTGGACTCCCTGGGTGAGCTTCCTA	875
Qy	841	AGCAGCTCGGATCTAGGCTAGCTTCAATGTGGAGCGGGCTTGTGACAGTACTTCCAC	900
Db	876	AGCAGCTCGGATCTAGGCTAGCTTCAATGTGGAGCGGGCTTGTGACAGTACTTCCAC	935
Qy	901	TGTGTATCTCGATCAGCGGTCTCTAGGCCTATATAGGAATAAATCTTTCGAGAGTTT	960
Db	936	TGTGTATCTCGATCAGCGGTCTCTAGGCCTATATAGGAATAAATCTTTCGAGAGTTT	995
Qy	961	CTTGAGGCTGTACACAGAGTTCTGGGTACCCACTGCTACTTTTGGGGATGACCGG	1020
Db	996	CTTGAGGCTGTACACAGAGTTCTGGGTACCCACTGCTACTTTTGGGGATGACCGG	1055
Qy	1021	CACCTCACCAACCGCATGCTCAGCATGGGTATGCTTACCAAGTA	1064
Db	1056	CACCTCACCAACCGCATGCTCAGCATGGGTATGCTTACCAAGTA	1099

BC035837 2119 bp mRNA linear PRI 30-JUN-2004
 Homo sapiens hyaluronan synthase 1, mRNA (cDNA clone MGC:46218
 IMAGE:589083), complete cds.

ACCESSION BC035837
 VERSION BC035837.1 GI:23243101
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 2119)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Caeavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., C., Rodrigues, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smaluis, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 2119)
 Strausberg, R.
 Direct Submission
 Submitted (31-JUL-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
COMMENT

NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgapbs@mail.nih.gov
 Tissue Procurement: Invitrogen
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: http://www.nisc.nih.gov/
 Contact: nisc_mgc@hgrl.nih.gov
 Akter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsugeon, C., Vogt, J.D., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINL at: http://image.lnl.gov
 Series: IRAK Plate: 79 Row: f Column: 24
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 4504338.

FEATURES

source

Location/Qualifiers
 1..2119
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:46218 IMAGE:589083"
 /tissue_type="Ovary, pooled from 3 adults"
 /clone_lib="NIH MGC_125"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 1..2119
 /gene="HAS1"
 /note="synonym: HAS"
 /db_xref="LocusID:3036"
 /db_xref="MIM:601463"
 42..1775
 /gene="HAS1"
 /codon_start=1
 /product="hyaluronan synthase 1"
 /protein_id="AAH3837.1"
 /db_xref="GI:23243102"
 /db_xref="LocusID:3036"
 /translation="MRQAPKPTPAARCSGLARRVLTIFAALLILGLMTWYAAAGVP
 LASDRYGLAFGLAFSLHVAQLPAYLEHRRVAAARGLDAAATVALTISA
 YQEDPAYLQCLAGARALLYPRARLVLMVVDGNRAEDLYMVDMEFADBDPATYV
 WDNYPHQPWEPAAGAVGAGAVREVEADPGRLAEALVTRRCVCAQVRGKGREVM
 YTAFKALGDSNVQVQCDSTRLDPMALLELVRLVDEDPVGVAGVGDVRLNPLDSHW
 SFLSLRYWAFNVVERACQSYFHCVCSIGGLYRNLLQQLFLEAWYNQKPLGTHCT
 FGDDRHMTNRLMSGYATKYTSRSCYSETSSFLRWLSQQTFRWSKFRWLYNALW
 WHRHWTMYEAVSGLPFFVAATVRLFPYAGRPWALLVLLCVQGVAAKAAFAAW
 LRGLCRMVLLSLYAPLMCGLLPAKFLALVTMNOGCTSGRRKLAANVPLPLALW
 ALLGLGLVRSVAHEARADWSGPRAAEAYHLAAGAYGVGVWVWMLTYWGVVRLC
 RRRTGGYRVQV"

gene

CDS

ORIGIN

Query Match 98.8%; Score 1051.8; DB 9; Length 2119;
 Best Local Similarity 99.8%; Pred. No. 1.2e-134;
 Matches 1053; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 10 CAGGACGGCCCAAGCCCACTCTCTCAGCCCGCCCTGCTCGGCTCGGCCGGGGTG 69
 Db 48 CAGGACGGCCCAAGCCCACTCTCTCAGCCCGCCCTGCTCGGCTCGGCCGGGGTG 107
 QY 70 CTGACCATCGCTTCGCCCTCTCATCTCGGCTCATGACCTGGGCTACGCCCGCGG 129
 Db 108 CTGACCATCGCTTCGCCCTCTCATCTCGGCTCATGACCTGGGCTACGCCCGCGG 167

QY 130 GTGCGCTGGCTCGATCGCTACGGCTCTCGGCTCTCGGCTCTACGGGGCTTCCTT 189
Db 168 GTGCGCTGGCTCGATCGCTACGGCTCTCGGCTCTCGGCTCTACGGGGCTTCCTT 227
QY 190 TCAGCGCACCTGGTGGCGCAGAGCTCTTCGGCTACCTGGAGCACCGGGGTGCGGGCG 249
Db 228 TCAGCGCACCTGGTGGCGCAGAGCTCTTCGGCTACCTGGAGCACCGGGGTGCGGGCG 287
QY 250 GGGCGCGGGGGCGCTGTGATGACAGCACCGCGCGAGTGTGGCGCTGACCACTTCGCGC 309
Db 288 GGGCGCGGGGGCGCTGTGATGACAGCACCGCGCGAGTGTGGCGCTGACCACTTCGCGC 347
QY 310 TACCAAGGAGACCGCGCTACCTGGCGCAGTGTGGCTGGCGCGCGCGCTGTGTATC 369
Db 348 TACCAAGGAGACCGCGCTACCTGGCGCAGTGTGGCTGGCGCGCGCGCTGTGTATC 407
QY 370 CGCGCGCGCGCTGGCTCTCATGTGTGTGATGACCAACCGCGCGAGACCTCTAC 429
Db 408 CGCGCGCGCGCTGGCTCTCATGTGTGTGATGACCAACCGCGCGAGACCTCTAC 467
QY 430 ATGTGTCGATGTTCCGAGAGTCTTCGCTGACGAGACCCCGCACGTACGTGTGGAC 489
Db 468 ATGTGTCGATGTTCCGAGAGTCTTCGCTGACGAGACCCCGCACGTACGTGTGGAC 527
QY 490 GCACTACCAACCGCGCTGGAGACCCCGCGCGGGCGGGCGGGTGGCGCGGAGCTAT 549
Db 528 GCACTACCAACCGCGCTGGAGACCCCGCGCGGGCGGGTGGCGCGGAGCTAT 587
QY 550 CGGAGGTGGAGCGGAGATCCTGGCGGCTGGCAGTGGAGCGCTGTGAGGACTCGC 609
Db 588 CGGAGGTGGAGCGGAGATCCTGGCGGCTGGCAGTGGAGCGCTGTGAGGACTCGT 647
QY 610 AGTCCGCTGTGCTGGCGAGCGCTGGCGCGGCAAGCGCGAGGTCACTACAGCCTTC 669
Db 648 AGTCCGCTGTGCTGGCGAGCGCTGGCGCGGCAAGCGCGAGGTCACTACAGCCTTC 707
QY 670 AAGCGCTCGGAGATTCGATGAGTACGTCGAGGTCTGTGACTCGGACACAAAGGTGGAC 729
Db 708 AAGCGCTCGGAGATTCGATGAGTACGTCGAGGTCTGTGACTCGGACACAAAGGTGGAC 767
QY 730 CCATGCGACTGCTGGAGCTCGTGGGGTACTGAGCAGAGGACCCCGGGTAGGGCTGT 789
Db 768 CCATGCGACTGCTGGAGCTCGTGGGGTACTGAGCAGAGGACCCCGGGTAGGGCTGT 827
QY 790 GTGGGAGCGTGGGATCTTAACCTCTGAGCTCTGGGCTAGCTTCTTAAGCAGCCTG 849
Db 828 GTGGGAGCGTGGGATCTTAACCTCTGAGCTCTGGGCTAGCTTCTTAAGCAGCCTG 887
QY 850 GATACCTGGGTAGCTTCAATGTGGAGCGGGCTGTGAGGCTACTTCCACTGTGTATCC 909
Db 888 GATACCTGGGTAGCTTCAATGTGGAGCGGGCTGTGAGGCTACTTCCACTGTGTATCC 947
QY 910 TGCATCAGCGGTCTCTAGGCTATATAGGAATAACCTCTTGAGCAGTCTTCTGAGGCC 969
Db 948 TGCATCAGCGGTCTCTAGGCTATATAGGAATAACCTCTTGAGCAGTCTTCTGAGGCC 1007
QY 970 TGGTACACAGAGTTCTGGGTACCACTGTACTTTTGGGATGACCGGCACTCACC 1029
Db 1008 TGGTACACAGAGTTCTGGGTACCACTGTACTTTTGGGATGACCGGCACTCACC 1067
QY 1030 AACCGCATGCTCAGCATGGTTATGCTACCAAGTA 1064
Db 1068 AACCGCATGCTCAGCATGGTTATGCTACCAAGTA 1102

RESULT 4
HUMHAS HUMHAS 2108 bp mRNA linear PRI 06-FEB-1999
LOCUS Homo sapiens mRNA for hyaluronan synthase, complete cds.
DEFINITION D84424
ACCESSION D84424
VERSION D84424.1 GI:1401033
KEYWORDS hHAS; hyaluronan synthase.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2108)
AUTHORS Itano.N. and Kinata.K.
TITLE Molecular cloning of human hyaluronan synthase
JOURNAL Biochem. Biophys. Res. Commun. 222 (3), 816-820 (1996)
MEDLINE 96244584
PUBMED 8651928
REFERENCE 2 (bases 1 to 2108)
AUTHORS Itano.N.
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2108)
AUTHORS Itano.N.
TITLE Direct Submission
JOURNAL Submitted (16-APR-1996) Naoki Itano, Aichi Medical University,
Institute for Molecular Science of Medicine; Nagakute, Aichi-gun,
Aichi 480-11, Japan (Tel:052-264-4811(ex.2087), Fax:0561-63-3532)
FEATURES
source
1..2108
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="brain"
/clone_lib="igtl1"
/dev_stage="fetal"
1..2108
/gene="hHAS"
149..1780
/gene="hHAS"
/codon_start=1
/product="hyaluronan synthase"
/protein_id="BAAL2351.1"
/translation="MTWAYAAGVPLASDRYGLLAGFLYGLAFSLVAHQAQSLFAYLEHR
RVAARAGPLDAATARSVALTISAYQEDPAVLROCLASARALLYPRARVRVLMVDGN
RAEDLYMVDMPREVEADEPATYVMDGNVHPWEPAAAGVAGAGAYREVEADPRLA
VEALVTRRCVCVAQRWGGKREVMYAFKALGSDVDYVQVCDSTRDLPMALLLRV
LDEPRVGAAGVDRIINPLDSWVSFLSRLYVAFNVRACQSYFHCVCISGPLGL
YRNLLQOFLFLEAMYNQFLGCTGFGDRHLTNRMLSMGYATKTSRSRCYSETPSSF
LRWLSQOTRWSKSYFRELWYALNLMHRRHAWMTYEAUVSGLPFFVFAATVLRLLPMQ
PWALLWLLCVQVALAKAFAAWLRGLRMLVLSVAPLYMCELLPAKFLALVPMQ
SWGTSGRRKLAAANYVPLPLALWALLGLGLVPSVAHEARADWSGFSRAEAHYLAA
GAGAYGVWAMLTLYWGVRRLCRRRTGGTGVQV"

Query Match 98.7%; Score 1051.6; DB 9; Length 2108;
Best Local Similarity 99.6%; Pred. No. 1.3e-134;
Matches 1054; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 7 CAGCAGCAGCGCCCAAGCCACTCTCTGCAGCCCGCGCTGCTCGGCTCGCGGAGG 66
Db 50 CAGCAGCAGCGCCCAAGCCACTCTCTGCAGCCCGCGCTGCTCGGCTCGCGGAGG 109
QY 67 GTGTGTACCATCGCTTCGCCCTCTCATCTCGGGCTCATGACTGGGCTACGCCGCC 126
Db 110 GTGTGTACCATCGCTTCGCCCTCTCATCTCGGGCTCATGACTGGGCTACGCCGCC 169
QY 127 GGGGTGCGCTGGCTCCGATCGCTACGGCTCTCGGGCTTCGGCTTCAGCGGCTTC 186
Db 170 GGGGTGCGCTGGCTCCGATCGCTACGGCTCTCGGGCTTCGGCTTCAGCGGCTTC 229
QY 187 CTTTCAGCGCACCTGTGTGGCGCAGAGCTCTTTCGGTACTCTGGAGACACCGCGCGGTGCG 246
Db 230 CTTTCAGCGCACCTGTGTGGCGCAGAGCTCTTTCGGTACTCTGGAGACACCGCGCGGTGCG 289
QY 247 GCGGCGCGCGGGCGCTGGATGACCCACCGCGCGAGTGTGGCTGACCATCTCC 306
Db 290 GCGGCGCGCGGGCGCTGGATGACCCACCGCGCGAGTGTGGCTGACCATCTCC 349
QY 307 GCCTACCAAGGAGACCCCGGCTACTCTGGCGCAGTGTCTGGCGTCCGCCCGCGCTGTG 366
Db 350 GCCTACCAAGGAGACCCCGGCTACTCTGGCGCAGTGTCTGGCGTCCGCCCGCGCTGTG 409

```

QY 367 TACCCGCGCGCGCTGCGCGTCTCTCATGTGGTGGATGGCAACCGCGCCGAGGACCTC 426
Db 410 TACCCGCGCGCGCGCTGCGCGTCTCTCATGTGGTGGATGGCAACCGCGCCGAGGACCTC 469
QY 427 TACATGGTGCATGTTTCGCGAGAGTCTTCGCTGACGAGGACCCCGCCACGTAAGTGTGG 486
Db 470 TACATGGTGCATGTTTCGCGAGAGTCTTCGCTGACGAGGACCCCGCCACGTAAGTGTGG 529
QY 487 GACGGCAACTACACAGCCCTGGGAACCGCGCGCGCGCGCGCTGCGCGCGCGCGAGCC 546
Db 530 GACGGCAACTACACAGCCCTGGGAACCGCGCGCGCGCGCGCTGCGCGCGCGCGAGCC 589
QY 547 TATCGGAGGTGAGGCGGAGGATCTTGGCGCGCTGGCAGTGGAGGCGCTGTGTGAGGACT 606
Db 590 TATCGGAGGTGAGGCGGAGGATCTTGGCGCGCTGGCAGTGGAGGCGCTGTGTGAGGACT 649
QY 607 CGCAGGTGCGTGTGCTGGCGCAGCGCTGGGGCGCGCGCGCGCGCGCGCTGATGTACAGCC 666
Db 650 CGCAGGTGCGTGTGCTGGCGCAGCGCTGGGGCGCGCGCGCGCGCGCTGATGTACAGCC 709
QY 667 TTCAAGGCGCTCGGAGATTCGCTGGAGTACGTCAGGCTCTGTGACTCGGACACAAAGTTG 726
Db 710 TTCAAGGCGCTCGGAGATTCGCTGGAGTACGTCAGGCTCTGTGACTCGGACACAAAGTTG 769
QY 727 GACCCCATGGCACTGCTGGAGCTCGTGGCGGTACTGGAGGAGACCCCGCGGTAGGGCT 786
Db 770 GACCCCATGGCACTGCTGGAGCTCGTGGCGGTACTGGAGGAGACCCCGCGGTAGGGCT 829
QY 787 GTTGTGGGAGCGTGGGATCTTAAACCTCTGGAGTCTGGAGTCTGGAGTCTGGAGTCTGGAG 846
Db 830 GTTGTGGGAGCGTGGGATCTTAAACCTCTGGAGTCTGGAGTCTGGAGTCTGGAGTCTGGAG 889
QY 847 CTGCGATCTGGGTAGCTTCAATGAGCGCGGCTTGTGAGGAGTCTGGAGTCTGGAGTCTGGAG 906
Db 890 CTGCGATCTGGGTAGCTTCAATGAGCGCGGCTTGTGAGGAGTCTGGAGTCTGGAGTCTGGAG 949
QY 907 TCTGATCAGCGGCTCTAGGCGCTATAGGAATAACCTCTTGCAGCAGTCTTCTTGGAG 966
Db 950 TCTGATCAGCGGCTCTAGGCGCTATAGGAATAACCTCTTGCAGCAGTCTTCTTGGAG 1009
QY 967 GCCTGGTACACAGAGTCTTGGGTACCCACTGTATCTTTGGGGATGACCGGACCTC 1026
Db 1010 GCCTGGTACACAGAGTCTTGGGTACCCACTGTATCTTTGGGGATGACCGGACCTC 1069
QY 1027 ACCAACCGCATCTCAGCATGGGTATGTACCAAGTA 1064
Db 1070 ACCAACCGCATCTCAGCATGGGTATGTACCAAGTA 1107

```

```

RESULT 5
ARL37038
LOCUS ARL37038 2117 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6162908.
ACCESSION ARL37038
VERSION ARL37038.1 GI:14478288
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 2117)
TITLES Itano,N. and Kimata,K.
POLYPEPTIDE Polypeptide of human-origin hyaluronate synthetase and DNA encoding
the same
JOURNAL Patent: US 6162908-A 1 19-DEC-2000;
FEATURES Location/Qualifiers
source 1..2117
/mol_type="unassigned DNA"
ORIGIN
Query Match 98.7%; Score 1051.6; DB 6; Length 2117;
Best Local Similarity 99.6%; Pred. No. 1.3e-134;

```

```

Matches 1054; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 7 CAGCAGGACGGCCCAAGCCCACTCTGACAGCCCGCGCTGCTCGGCTGCGCCGAGG 66
Db 50 CAGCAGGACGGCCCAAGCCCACTCTGACAGCCCGCGCTGCTCGGCTGCGCCGAGG 109
QY 67 GTGCTGACCATCGCTTCCGCTGCTCATCTCTGGGCTCATGACCTGGGCTGAGCCGCT 126
Db 110 GTGCTGACCATCGCTTCCGCTGCTCATCTCTGGGCTCATGACCTGGGCTGAGCCGCT 169
QY 127 GGGGTGCGCTGCGCTTCCGCTGCTCATGAGCCCTCTTGGGCTTGGGCTCTACCGGGGCTTC 186
Db 170 GGGGTGCGCTGCGCTTCCGCTGCTCATGAGCCCTCTTGGGCTTGGGCTCTACCGGGGCTTC 229
QY 187 CTTTTCAGCGCAGCTGGTGGCGCAGAGCTCTTTCGGGTACCTTGGAGCACCGGGCGGTGGGG 246
Db 230 CTTTTCAGCGCAGCTGGTGGCGCAGAGCTCTTTCGGGTACCTTGGAGCACCGGGCGGTGGGG 289
QY 247 GCGCGCGCGCGGGGCGCTGGATGCAACCCCGCGCAGTGTGGCGCTGACCATCTCC 306
Db 290 GCGCGCGCGCGGGGCGCTGGATGCAACCCCGCGCAGTGTGGCGCTGACCATCTCC 349
QY 307 GCCTACAGGAGGAGCCCGCGCTACCTGCGCAGTGTGGCGCTGCGCGCGCGCTGCTG 366
Db 350 GCCTACAGGAGGAGCCCGCGCTACCTGCGCAGTGTGGCGCTGCGCGCGCGCTGCTG 409
QY 367 TACCCGCGCGCGCTGCGGCTCTCATGTGGTGGATGGCAACCGCGCGCGAGGACCTC 426
Db 410 TACCCGCGCGCGCTGCGGCTCTCATGTGGTGGATGGCAACCGCGCGCGAGGACCTC 469
QY 427 TACATGTGCGATGTTTCCGCGAGGTCTTCCGTGACGAGGACCCCGCCACGTAAGTGTGG 486
Db 470 TACATGTGCGATGTTTCCGCGAGGTCTTCCGTGACGAGGACCCCGCCACGTAAGTGTGG 529
QY 487 GACGGCAACTACACAGCCCTGGGAACCGCGCGCGCGCGCGCTGAGGAGTCTTCTTGGAG 546
Db 530 GACGGCAACTACACAGCCCTGGGAACCGCGCGCGCGCGCGCTGAGGAGTCTTCTTGGAG 589
QY 547 TATCGGAGGTGAGGCGGAGGATCTTGGCGGCTTGGAGTCTGGAGTCTGGAGTCTGGAGTCT 606
Db 590 TATCGGAGGTGAGGCGGAGGATCTTGGCGGCTTGGAGTCTGGAGTCTGGAGTCTGGAGTCT 649
QY 607 CGCAGGTGCGTGTGCTGGCGCAGCGCTGGGGCGCGCAAGCGCGAGTCTATGTACAGCC 666
Db 650 CGCAGGTGCGTGTGCTGGCGCAGCGCTGGGGCGCGCAAGCGCGAGTCTATGTACAGCC 709
QY 667 TTCAAGGCGCTCGGAGATTCGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 726
Db 710 TTCAAGGCGCTCGGAGATTCGCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG 769
QY 727 GACCCCATGGCACTGCTGGAGCTCGTGGCGGTACTGGAGGAGACCCCGCGGTAGGGCT 786
Db 770 GACCCCATGGCACTGCTGGAGCTCGTGGCGGTACTGGAGGAGACCCCGCGGTAGGGCT 829
QY 787 GTTGTGGGAGCGTGGGATCTTAAACCTCTTGGACTCTTGGGCTGAGTCTTCTTAAAGCAGC 846
Db 830 GTTGTGGGAGTGTGGGATCTTAAACCTCTTGGACTCTTGGGCTGAGTCTTCTTAAAGCAGC 889
QY 847 CTGCGATCTGGGTAGCTTCAATGAGCGCGGCTTGTGAGGAGTCTGGAGTCTGGAGTCTGGAG 906
Db 890 CTGCGATCTGGGTAGCTTCAATGAGCGCGGCTTGTGAGGAGTCTGGAGTCTGGAGTCTGGAG 949
QY 907 TCCTGATCAGCGGCTCTAGGCGCTATAGGAATAACCTCTTGCAGCAGTCTTCTTGGAG 966
Db 950 TCCTGATCAGCGGCTCTAGGCGCTATAGGAATAACCTCTTGCAGCAGTCTTCTTGGAG 1009
QY 967 GCCTGGTACACAGAGTCTTGGGTACCCACTGTATCTTTGGGGATGACCGGACCTC 1026
Db 1010 GCCTGGTACACAGAGTCTTGGGTACCCACTGTATCTTTGGGGATGACCGGACCTC 1069
QY 1027 ACCAACCGCATCTCAGCATGGGTATGTACCAAGTA 1064
Db 1070 ACCAACCGCATCTCAGCATGGGTATGTACCAAGTA 1107

```

```
RESULT 6
AY463695
LOCUS
DEFINITION
ACCESSION
VERSION
AY463695.1 GI:38607341
KEYWORDS
SOURCE
ORGANISM
Papio anubis (olive baboon)
Papio anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopithecinae; Papio.
REFERENCE
1 (bases 1 to 2109)
Martinez-Duncker,I., Oriol,R. and Mollicone,R.
Evolution of the hyaluronan, nodulation c, chitin and cellulose
synthases: a superfamily of cell-wall associated carbohydrate
polymerizing enzymes
Unpublished
JOURNAL
2 (bases 1 to 2109)
Martinez-Duncker,I., Oriol,R. and Mollicone,R.
Direct Submission
AUTHORS
TITLE
Submitted (12-NOV-2003) Unite 504 Glycobiology Et Signalisation
Cellulaire, Inserm, 16 Avenue Paul Vaillant Couturier, Villejuif
94807, France
FEATURES
Location/Qualifiers
1..2109
/organism="Papio anubis"
/mol_type="mRNA"
/db_xref="taxon:9555"
1..2109
/gene="HAS1"
36..1787
/gene="HAS1"
/note="cell wall associated protein; produces hyaluronic
acid chains"
/codon_start=1
/product="hyaluronan synthase"
/protein_id="AAR2554.1"
/db_xref="GI:38607342"
/translation="MTQRDTPKPTPAARRCSGLARRVLTIFAALLILGLMTWAYAAGV
PLTSDRYGLAFGLYGFALSARALLIYPRARLVLMVDGNRPEDLYMDMPREVFADE
DPATVMDGNVHQPHEPAANGVAGVREAEADPGELAVEALVTRCYCVAORWG
KREVMYTFALGDSVDYVQVDSDDLDPALLELVQVLEDDPRVAGVGDVRLIN
PLDSWSFLSLRYVWVAFNVRACQSYFHCVSCISGPIGLYRNLLQOFLQWYNQKF
LGTHCTFGDDRHLNRMLSMGYATKYTSRCSYSETPSSFLRWLSQQTRWSKSYFREW
LYNALMWRHHAMWTEYAVSGLFFPFFVAATVRLIFYAGRPAWLLVLLCVQGVALLAK
AAPAAWLRGLRMVLLSLYAPLYMCGLLPAKFLALVTMNSQSGWTSGRRKLAANYVPL
LPLALWALLLGLVRSVAHEARADWSGPRAAEAYHLAAGAYGVGVVWMLTYWV
GVRUCRRRTGGYRVQV"
ORIGIN
Query Match 92.6%; Score 986.2; DB 9; Length 2109;
Best Local Similarity 95.6%; Pred. No. 1,1e-125;
Matches 1031; Conservative 0; Mismatches 33; Indels 15; Gaps 1;

QY 1 ATGACACAGCAGGACGCCGCCAAGCCCACTCTCTGCAGCGCCGCCGCTGCTCGGCCCTGCC 60
Db |
36 ATGACACAGCGGACACGCCCAAGCCCACTCTCTGCAGCGCCGCCGCTGCTCGGCCCTGCC 95
QY 61 CGGAGGTCCTGACCATCGCTTCGCCCTTGCCTCATCTCGGGCTCATGACCTGGCCTAC 120
Db |
96 CGGAGGTCCTGACCATCGCTTCGCCCTTGCCTCATCTCGGGCTCATGACCTGGCCTAC 155
QY 121 GCCCGCGGGTGGCCGCTGCCATCGCTACGGCCCTCTGGCCTCTCGGCTCTACCGG 180
Db |
156 GCCCGCGGGTGGCCGCTGCCATCGCTACGGCCCTCTGGCCTCTCGGCTCTACCGG 215
QY 181 GCCTTCCTTTACAGCGCACTGGTGGCGCAGAGCCTCTTCCTGGGTACCTGGAGACCGCGG 240
Db |
216 GCCTTCCTCTCGGCGCACCTGTGGCGCAGAGCCTCTCTCGGTACCTGGAGCATCGCGG 275
```

```
QY 241 GTGGCGCGCGCGCG-----CGGGCGCGCTGGATGACAGCCACCGCGCGC 285
Db |
276 GTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGGATGACAGCCACCGCGCGC 335
QY 286 AGTGTGGCGCTGACCATCTCCGCCCTACACAGAGAGACCCCGCGGTACTCTCGCGCAGTGCCTG 345
Db |
336 AGCGTGGCGCTGACCAITTTCCGCCCTACACAGAGAGACCCCGCGGTACTCTCGCGCAGTGCCTG 395
QY 346 CGCTCGCGCGCGCGCGCTGCTGTATACCGCGCGCGCGCGCTGCGCGTCTCTCATGGTGGTGGAT 405
Db |
396 GTGTTCGCGCGCGCGCGCTGCTGTATACCGCGCGCGCGCGCTGCGCGTCTCTCATGGTGGTGGAC 455
QY 406 GGCAACCGCGCGCGAGGACCTCTACATGTGTGCACATGTTCCCGAGGTCTTCGCTGACCGAG 465
Db |
456 GGCAACCGCGCGCGAGGACCTCTACATGTGTAGACATGTTCCCGAGGTCTTCGCGCAGCGAG 515
QY 466 GACCCCGCGCAGTACTGTTGGGACGGCAACTACACACAGCCCTCGGAAACCCCGCGCGCGCG 525
Db |
516 GACCCCGCGCAGTACTGTTGGGACGGCAACTACACACAGCCCTCGGAAACCCCGCGCGCGGTG 575
QY 526 GGCGCGGTGGCGCGCGGAGCCCTATCGGAGGTGAGAGCGGAGGATCTCTGGCGCGCTGGCA 585
Db |
576 GGCGCGGTGGCGCGCTCGGAGCCCTACCGGAGGTGGAGCGGAGGATCCCGCGCGGTGGCG 635
QY 586 GTGGAGCGCGTGGTGAGGACTCGCAGGTGCGTGTGCGTGGCGCAGCGCTGGCGCGGCAAG 645
Db |
636 GTGGAGCGCGTGGTGAGGACTCGCAGGTGCGTGTGCGTGGCGCAGCGCTGGCGCGGCAAG 695
QY 646 CGCGAGGTCTATGTATACACAGCCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGTTC 705
Db |
696 CGCGAGGTCTATGTATACACCGCCCTTCAAGCGCTCGGAGACTCGGTGGACTACGTGCAAGTTC 755
QY 706 TGTGACTCGGACACAAGGTTGGACCCCATGCGCACTGCTGAGAGCTCGTGGCGGTACTGGAC 765
Db |
756 TGTGACTCGGACACAAGGTTGGACCCCATGCGCACTGCTGAGAGCTCGTGGCGGTACTGGAT 815
QY 766 GAGCACCCCGGGTAGGGCTGTTGTGGGACGCTGGGATCCTTAAACCCCTCGGACTCC 825
Db |
816 GAGCACCCCGGGTAGGGCTGTTGTGGGACGCTGGGATCCTTAAACCCCTCGGACTCC 875
QY 826 TGGGTGAGCTTCTTAAGCAGCCTCGGATACCTGGGTAGCCTTCAATGTGGAGCGCGCTTGT 885
Db |
876 TGGGTGAGCTTCTTAAGCAGCCTCGGATACCTGGGTAGCCTTCAATGTGGAGCGCGCTTGT 935
QY 886 CAGAGCTACTTCCACTGTGTATCTCGCATCAGCGGTCTCTAGGCGCTATATAGGAATAAC 945
Db |
936 CAGAGCTACTTCCACTGTGTATCTCGCATCAGCGGTCTCTAGGCGCTATATAGGAATAAC 995
QY 946 CTCTTGACGACGTTTCTTGAGGCGCTGTACAAACAGAAAGTTCTCTGGGTACCCACTGTACT 1005
Db |
996 CTCTTGACGACGTTTCTTGAGGCGCTGTGTACAAACAGAAAGTTCTCTGGGNAACCCACTGTACT 1055
QY 1006 TTTGGGATGACCGCGCACCTCACCAACCGCATGCTCAGCATGGGTATGTCTACCAAGTA 1064
Db |
1056 TTTGGGATGACCGCGCACCTCACCAACCGCATGCTCAGCATGGGTATGTCTACCAAGTA 1114
RESULT 7
AB097568 2095 bp mRNA linear ROD 28-APR-2004
LOCUS
DEFINITION
Rattus norvegicus HAS1 mRNA for hyaluronan synthase 1, complete
cde.
AB097568
VERSION
AB097568.1 GI:26453348
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1
AUTHORS
Itano,N., Sawai,T., Atsumi,F., Miyaishi,O., Taniguchi,S.,
```

Kannagi, R., Hamaguchi, M. and Kimata, K.
 Selective expression and functional characteristics of three
 Mammalian hyaluronan synthases in oncogenic malignant
 transformation
 J. Biol. Chem. 279 (18), 18679-18687 (2004)
 14724275
 2 (bases 1 to 2095)
 Direct Submission
 Submitted (05-DEC-2002) Naoki Itano, Aichi Medical University,
 Institute for Molecular Science of Medicine; Nagakute, Aichi, Aichi
 480-1195, Japan (E-mail: itano@amugw.aichi-med-u.ac.jp,
 Tel: 81-52-264-4811 (ex.2095), Fax: 81-561-63-3532)

FEATURES

Location/Qualifiers
 1..2095
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 1..2095
 43..1794
 /gene="HAS1"
 /gene="HAS1"
 /codon_start=1
 /product="hyaluronan synthase 1"
 /protein_id="BAC43730.1"
 /db_xref="GI:26453349"
 /translation="NRQDMPKPESEAAACGSLARRVLTITFALLILGLMTWYAAAGVP
 LASDPVGLAFLYGLAFILVAHVAOSLHVALEHRRVTVAAARFAKPLDAAATARSVA
 LTIAYQEDPTVQLCLISARALLYPRFLRLVLMVVDGNRAEDLWYDFRFRVAFED
 PATYVNDGNYHQPFAEAGAGVGEAEVREAEVDFRLAVALVLRRCVCAORWG
 GREVMYTAFAKAGSDVYQVDSDFDLPMALLEVRVLEDFRVAAGVGDVRLIN
 PLDSVFLSSLYRVAFNERASQYFHCVSCI SGPGLYRNLLQDFLEAWYNQKF
 LGTHCTFGDRHLTNRMLSMGVATKYSRSCYSETSPSLFLWLLQOTRWSKSYFREW
 LYNALWHRHRAWMTVEAVVSGLPFPFVAATVRLFYAGRPWALLWLLLCVQVALAK
 AFAFWLRCLRMVLLSLYAPLYMGLPLPAKFLALVTWNSGWTSGRKKLAANTVPV
 LPLWALLLLGLLRSAVAQEVRAVDSGPRAAEYHLLAAGASAYVAVVWIMLTIIYW
 GVRRLRRSSGGYRVQV"

ORIGIN

Query Match 74.3%; Score 791.4; DB 10; Length 2095;
 Best Local Similarity 85.2%; Pred. No. 4.7e-99;
 Matches 914; Conservative 0; Mismatches 141; Indels 18; Gaps 2;

QY 10 CAGGAGCGCCCAAGCCCACTCTCTGAGCCCGCGCTGCTCGGCGCTGCGCGCGAGGGTG 69
 |||||
 DB 49 CAGGACATGCCAAGCCCTCAGAGCAGCACGCTGCTGCTCGGCTGCGCAGCGGGTG 108
 |||||
 QY 70 CTGACCATCGCTTGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCGCGCGG 129
 |||||
 DB 109 CTCAGCATCACCTTGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCGCAGGA 168
 |||||
 QY 130 GTGCGCTGCTCGCATCGCTACGCGCTCTCGGCTCTGCGGCTCTACGCGGCTTCTTT 189
 |||||
 DB 169 GTACCTCTGGCTTGTATCCCTATGCGCTCTGCGCTTGGGCTCTATGGGCGTTCCTC 228
 |||||
 QY 190 TCAGCGCACCTGGTGGCGCAGAGCTCTTCGGGTACTTGGAGCAGCGCGGTGGCGGG 249
 |||||
 DB 229 AGTGGCACCTAGTGGCAGAGGCTCTTCGCTTACCTGGAGCAGCGAGGTGACCGTG 288
 |||||
 QY 250 GCGGCGCG-----GGGCGCGTGGATGACGACCAACCGCGCGCATGTGGCGG 294
 |||||
 DB 289 GCTGGCGCGCGCGCTTTTTCGAAGGAGACCCCTGGATGCGGCGCTGCGCGCGGTGGCA 348
 |||||
 QY 295 CTGACCATCTCGCTTACGAGGAGNACCGCGGTACTTGGCGCATGCTGCGCTCGCGC 354
 |||||
 DB 349 CTCACCATCTCAGCTTACGAGGAGACCCCACTTACCTGCGCGCATGTGACCTCGCGG 408
 |||||
 QY 355 CGCGCGCTCTGTACCGCGCGCGCGGTGCGGCTCTCATGCTGTGTGATGGCAACCGC 414
 |||||
 DB 409 CGCGCTCTGTGTACCGCGCGCGCGGTGCGGCTGTATGTTGTAGACCGCATCGC 468
 |||||
 QY 415 GCCGAGGACCTTACATGTTGACATGTTCCGCGAGGCTTCTCGCTGACGAGGACCCCGC 474
 |||||

Db 469 GCGGAGGATCTGTATCATGGTGGACATGTTCCGAGAAAGTCTTCGACAGATGAGGACCTGCC 528
 QY 475 ACGTACGTGTGGGACGCGCAATACCAACAGCCCTCGGAAACCCGC---GCGCGCGCGCG 531
 |||||
 Db 529 ACTTATGTGTGGATGGCACTTACCATTAGCTTGGGAGCCAGCTGAGGCGCGGGTGTCT 588
 |||||
 QY 532 GTGGCGCGGAGCCATTATCGGAGGTGGAGGCGGAGGATCTTGGCGGCTGGCAGTGGAG 591
 |||||
 Db 589 GTGGGTGAAGGTGCCCTACCGGAGGTGGAGGCTGAGGACCCCTGGCGCGCTGGCGGTAGAG 648
 |||||
 QY 592 GCGCTGTGGAGACTCGCAGGTGCTGCTGGCGGAGCGCTGGCGGCGGCGGCGGAGCGGAG 651
 |||||
 Db 649 GCGCTGTGGAGGAGCCCGAGGTGCTGCTGGGCTGAGCGCTGGGGTGGCAAGCGCGAA 708
 |||||
 QY 652 GTCATGTACACAGCCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCGAGTCTGTGAC 711
 |||||
 Db 709 GTCATGTACAGGCTTTCAGGCACTGGGTGACTCCGTGGACTACGTGCGAGGTCTGTGAC 768
 |||||
 QY 712 TCGGACACAAGTTGGACCCCATGTCATGCTGGAGTCTGTCGGGTACTTGGACGAGGAC 771
 |||||
 Db 769 TCAGACACAAGTTAGACCCCATGTCATGCTGGAGCTTGTCCGAGTGTCTGGATGAAGAC 828
 |||||
 QY 772 CCCCGGTAGGGCTGTTGTTGGGACCGTGGGATCCTTAACCTCTGCACTCCTGGGTC 831
 |||||
 Db 829 CCCCGGTAGGTGCTTGGAGGAGATGTGAGAAATCCTTAACCTCTGCACTCCTGGGTC 888
 |||||
 QY 832 AGCTTCTTAAGCAGCTCGATATCTGGGTAGCTTCAATGTGGAGCGGCTTGTGACAGGC 891
 |||||
 Db 889 AGCTTCTTGGAGGCTTTCGATATCTGGGTAGCTTCAATGTGGAGCGGCTTGTGACAGGC 948
 |||||
 QY 892 TACTTCCATCTGTATCTTCGATCAGCGGTCTCTTAGGCTTATATAGGAATAACCTCTTG 951
 |||||
 Db 949 TACTTCACTGTGTCTCTGTCATCAGTGTCTCTGGGTCTATACAGAAAACAATCTCTG 1008
 |||||
 QY 952 CAGCAGTCTTGTAGGCTGTGACAAACAGAAATCTCTGGGTACCCACTGACTTCTGGG 1011
 |||||
 Db 1009 CAGCAGTCTTGTGGAGGCTGTGTACAAACAGAAATCTCTGGGCACTTGTGCAATTTGG 1068
 |||||
 QY 1012 GATGACCGGACCTCACCACCGCATGCTCAGCATGGTGTATGCTACCAAGTA 1064
 |||||
 Db 1069 GATGACGACGACCTCACCACCGCATGCTTACGATGGGCTACGCTACCAAGTA 1121
 |||||
 RESULT 8
 E13681
 LOCUS
 DEFINITION DNA encoding novel mouse hyaluronate synthetase.
 ACCESSION E13681
 VERSION E13681.1 GI:3252450
 KEYWORDS JP 1997224674-A/1.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Itano, N. and Kimata, H.
 POLYPEPTIDE OF NEW HYALURONIC ACID-SYNTHETIC ENZYME AND DNA CODING
 THE SAME
 Patent: JP 1997224674-A 1 02-SEP-1997;
 REFERENCE 1 (bases 1 to 2102)
 AUTHORS Itano, N. and Kimata, H.
 TITLE POLYPEPTIDE OF NEW HYALURONIC ACID-SYNTHETIC ENZYME AND DNA CODING
 JOURNAL THE SAME
 COMMENT Patent: JP 1997224674-A 1 02-SEP-1997;
 OS Mus musculus (mouse)
 PN JP 1997224674-A/1
 PD 02-SEP-1997
 PF 26-FEB-1996 JP 1996038336
 PI ITANO NAKKI, KIMATA HIROHARU
 PC C12N1/09, C07K14/04, C07K14/47, C12N9/00, C12N1/21, C12R1/19; CC
 (C12N9/00,
 PC C12R1/91), (C12N9/00, C12R1/19), (C12N1/21, C12R1/19); CC
 strandedness: Double;
 CC topology: Linear;
 CC hypothetical: No;
 CC anti-sense: No;
 FH Key Location/Qualifiers


```
QY 250 GCGGCGCG-----GGGGCGCTGGATGACAGCAACCGCGCGCAGTGTGGCG 294
Db 295 GCTGCGCGCGCTCTTGGCGAAGGGCGCCCTGGATGCGGCACATGACGACGCGTGGCA 354
QY 295 CTGACCATCTCGCTACAGAGGACCCGCGTACTTGGCCAGTGCCTGGCGTCCGC 354
Db 355 CTCACCATCTGACCTTACCAAGAGGATCCCGTTACCTGCGCGAGTGTGACCTCCGC 414
QY 355 CGCGCCCTCTGTACCCCGCGCGCGCTGCGCTCTCATGTTGGTGGATGCAACCGC 414
Db 415 CGCGCTTGTCTACCCGACACGAGGTTACGCGTGTCTATGTTGGTGGACGCAACCGC 474
QY 415 GCGGAGGACCTTACATGTTGACATGTTCCGCGAGTCTTTCGTGACGAGACCCCGC 474
Db 475 GCTGAGGATCTGTACATGTTGACATGTTCCGAGAGTCTTCCCGCATGAGGACCCCGC 534
QY 475 AGCTAGTGTGGAGCGCACTACACAGAGCTTGGAGACCCCGCGG---CGGCGCG 531
Db 535 ACTTATGTTGGATGGCACTACCATGACCCCTGGAAACCAAGCGAGGCTACGCGCGT 594
QY 532 GTGGGCGCGGACCTTATCGGAGGTGGAGGGAGGATCCTGGGCGGTGCGAGTGGAG 591
Db 595 GTCGTTGAGGTGCTTACCGGAGGTGGAGGGAGGACCCCGGCGGTGGCGTGGAG 654
QY 592 GCGCTGTTGAGGACTCGCAGGTGCGTGTGCGTGGCGCAGCGTGGGCGGCAAGCGCGAG 651
Db 655 GCGCTGTTGAGAAACACGCAAGTGTGCGTGGCTGACGCGTGGGCGGCAACGTTGAG 714
QY 652 GTCATGTATACAGACCTTCAAGCGCTCGGAGATTCGTTGGAGCTACGTTGAGTCTGTGAC 711
Db 715 GTCATGTATACAGACCTTCAAGCGCTCGGCGACTCCGTTGGACTACGTTGAGTCTGTGAC 774
QY 712 TCGGACACAGGTGTGACCCCATGGCACTGCGAGCTGCGGGTACTGACAGGAGAC 771
Db 775 TCAGACACAGACTAGACCCCATGGCACTGCTGGAGCTTGTGGAGTGTGGATGAAGAC 834
QY 772 CCGCGGTAGGGGTGTTGGTGGGACGTGCGGATCCTTAAACCTCTGGAACCTCTGGGTC 831
Db 835 CCGCGGTAGGGGTGTTGGAGGGGATGTGAGGATCCTTAAACCTCTGGAACCTCTGGGTC 894
QY 832 AGCTTCTTAAAGCAGCTGCGATCTGGGTAGCTTCAATGTGGAGGGCGCTGTGACAGC 891
Db 895 AGCTTCTTGAAGCAGCTTCTGATCTGGGTAGCTTCAATGTGGAGGGCGCTGTGACAGC 954
QY 892 TACTTCCACTGTATCTTCCATCAGCGGTCTTAGGCTATATAGGATTAACCTCTG 951
Db 955 TACTTCCACTGTATCTTCCATCAGCGGTCTTAGGCTATATAGGATTAACCTCTG 1014
QY 952 CAGCAGTCTTGTAGGCGCTGTACACAGAGGTTCTTGGGTACCCACTGTACTTTTGG 1011
Db 1015 CAGCAGTCTTGTAGGCGCTGTACACAGAGGTTCTTGGGTACCCACTGTACTTTTGG 1074
QY 1012 GATGACCGGACCTCACCACCGCATGTCTCAGCATGGGTATATGCTACCAAGTA 1064
Db 1075 GATGACCGGACCTCACCACCGCATGTCTCAGCATGGGTATATGCTACCAAGTA 1127
```

```
RESULT 10
E30971
LOCUS 2102 bp DNA linear PAT 18-JUN-2001
DEFINITION Hyaluronic acid synthetase modified protein.
ACCESSION E30971
VERSION E30971.1 GI:13017286
KEYWORDS JP 2000004886-A/1.
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2102)
Naoki, I., Mamoru, Y. and Koji, K.
Hyaluronic acid synthetase modified protein
Patent: JP 2000004886-A 1 11-JAN-2000;
```

```
SEIKAGAKU KOGYO CO LTD
Mus sp. (mouse)
JP 2000004886-A/1
11-JAN-2000
24-JUN-1998 JP 1998193788
NAOKI ITANO, MAMORU YOSHIDA, KOJI KIMATA
C12N15/09, C12N9/00, C12N15/00
Strandedness: Double;
Topology: Linear;
Key Location/Qualifiers
FT CDS 49..1800.
FEATURES
source
1..2102
/organism="Mus sp."
/mol_type="genomic DNA"
/db_xref="taxon:10095"
ORIGIN
Query Match 73.9%; Score 786.6; DB 6; Length 2102;
Best Local Similarity 84.9%; Pred. No. 2.1e-98;
Matches 911; Conservative 0; Mismatches 144; Indels 18; Gaps 2;
QY 10 CAGGACGCGCCCAAGCCCACTCTCTGACGCGCGCGCTGCTCCGCGCTGGCCCGGAGGTG 69
Db 55 CAGGACATGCCCAAGCCCTCAGAGCAGCGCGTGTCTGCTCTGGCTTGGCCGCGAGCA 114
QY 70 CTGACCATCGCTTCGCGCTGCTCATCTGGGCTCATGACCTGGGCTTACGCGCGCGG 129
Db 115 CTCAGATCATCTTTGGCTGCTCATCTGGGCTCATGACCTGGGCTTACGCGCGAGC 174
QY 130 GTGCGCTGGGCTCCGATCGCTACGCGCTCTCTGGCTTTCGGGCTCTACGGGCTTCTT 189
Db 175 GTTCTCTGGCTTCAGATCGCTATGACTCTCTGGCTTTCGGCTCTATGGGCACTTCTC 234
QY 190 TCAGGCACTCTGGTGGCGCAGAGCTCTTGGGTACCTTGGAGCACCGCGCGGTGGCGG 249
Db 235 AGCGCACACCTAGTGGCACAGAGCTCTTGGCTTACCTGGAGCACCGAAGGTTGGCAGC 294
QY 250 GCGGCGCG-----GGGCGCGTGGATGACGACCCCGCGCAGTGTGGCG 294
Db 295 GCTGCGCGCGCTCTTGGCGAAGGGGCGCTTGGATGGCGCACTGACAGCGTGGCA 354
QY 295 GTGACCATCTCGCTTACAGAGGACCCCGCTACCTGCGCGAGTGTGCTGGCGTCCGCG 354
Db 355 CTCACCATCTCAGCTTACCAAGAGATCCCGCTTACCTGGCGAGTGTGACCTCCGCG 414
QY 355 CGCGCTTGTGTACCCCGCGCGCTGCGCTCTCATGTTGGTGGATGGAACCGC 414
Db 415 CGCGCTTGTGTACCCCGCACACGAGGTTACGCGTCTCATGTTGGTGGACGCGCAACCGC 474
QY 415 GCGGAGGACCTTACATGTTGACATGTTCCGCGAGGTTCTTGGGTGACGAGGACCCCGC 474
Db 475 GCTGAGGATCTGTATCATGTTGGACATGTTCCGAGAGTCTTTCGCGATGAGACCCCGC 534
QY 475 AGCTAGTGTGGAGCGCAACTACCAAGCGCTGGGAAACCGCGCGCG---CGGCGCG 531
Db 535 ACTTATGTTGGATGGCACTTACCATGACCCCTGGAAACCAAGCGAGGCTACGCGCGT 594
QY 532 GTGGCGCGGAGCTTATCGGAGGTGGAGCGGAGGATCTTGGCGGTGCGAGTGGAG 591
Db 595 GTCGTTGAGGTGCTTACCGGAGGTGGAGGGAGGACCCCGGCGGTGGCGTGGAG 654
QY 592 GCGCTGTTGAGGACTCGCAGGTGCGTGTGCGTGGCGCAGCTGGGCGGCAAGCGCGAG 651
Db 655 GCGCTGTTGAGAAACACGCAAGTGTGCGTGGCTGACGCGTGGGCGGCAACGTTGAG 714
QY 652 GTCATGTATACAGACCTTCAAGCGCTCGGAGATTCGTTGGAGCTACGTTGAGTCTGTGAC 711
Db 715 GTCATGTATACAGCTTTCAGGCACTGGCGAGTCTCCGTTGGACTACGTTGAGTCTGTGAC 774
QY 712 TCGGACACAGGTGTGACCCCATGGCACTGCGAGCTGCGGGTACTGACAGGAGAC 771
```

```
REFERENCE
AUTHORS Naoki, I., Mamoru, Y. and Koji, K.
TITLE Hyaluronic acid synthetase modified protein
JOURNAL Patent: JP 2000004886-A 1 11-JAN-2000;
```


Db 775 TCAGACACAGACTAGACCCCATGGCACTGCTGGAGCTGTGGAGTGTGGATGAAGAC 834
QY 772 CCCCCGGTAGGGCTGTGGTGGGAGCGTGGGATCCCTTAACCTCTCTGGAATCTCTGGGTC 831
Db 835 CCCCCGGTAGGGCTGTGGTGGGAGCGTGGGATCCCTTAACCTCTCTGGAATCTCTGGGTC 894
QY 832 AGCTTCTTAAGCAGCTGCGATCTGAGTACCTGGGTAGCTTCAATGTGGAGCGGGCTGTGCAGAGC 891
Db 895 AGCTTCTTAAGCAGCTGCGATCTGAGTACCTGGGTAGCTTCAATGTGGAGCGGGCTGTGCAGAGC 954
QY 892 TACTTCCACTGTGTATCTCTGCATCAGCGGTCTCTAGGCCTATATAGGAATAACCTCTTG 951
Db 955 TACTTCCACTGTGTATCTCTGCATCAGTGTCTCTGGGTCTATACAGAAAATCTCTCTG 1014
QY 952 CAGCAGTTCTTGAAGCGCTGTGATCAACACAGAAATCTCTGGGTACCCACTGTACTTTTGGG 1011
Db 1015 CAGCAGTTCTTGGAGCGCTGTGATCAACACAAAGTTCTCTGGGCACCACTGCACATTTGGG 1074
QY 1012 GATGACCGGCACCTCACCACCGCATGCTCAGCATGGGTATGTACCAAGTA 1064
Db 1075 GATGACCGGCACCTCACCACCGCATGCTTAGCATGGGTATGTACCAAGTA 1127

RESULT 11
E34326 E34326 2102 bp DNA linear PAT 31-JAN-2002
LOCUS DNA for gene targeting hyaluronic acid synthase gene.

DEFINITION E34326
ACCESSION E34326
VERSION E34326.1 GI:18624311
KEYWORDS JP 2000116382-A/1.
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE Itano,N. and Kimata,K.
AUTHORS 1 (bases 1 to 2102)
TITLE DNA for gene targeting hyaluronic acid synthase gene
JOURNAL Patent: JP 2000116382-A 1 25-APR-2000;
SEIKAGAKU KOGYO CO LTD

COMMENT OS Mus musculus (mouse)
PN JP 2000116382-A/1
PD 25-APR-2000
PF 13-OCT-1998 JP 1998291201

PR NAOKI ITANO,KOJI KIMATA
PC C12N15/09,C12N5/10,G01N33/50//C12N9/00,C12Q1/68,(C12N15/09,PC
C12R1:91),
PC (C12N5/10,C12R1:91),C12N15/00,C12N5/00,(C12N15/00,C12R1:91),
PC (C12N5/00,C12R1:91)
CC

FEATURES
FT CDS Location/Qualifiers
49..1800.
1..2102
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

ORIGIN
Query Match 73.9%; Score 786.6; DB 6; Length 2102;
Best Local Similarity 84.9%; Pred. No. 2.1e-98;
Matches 911; Conservative 0; Mismatches 144; Indels 18; Gaps 2;

QY 10 CAGACGCGCCCAAGCCCACTCTCTCAGCGCGCGCTCTCGGCTTGGCGGAGGTG 69
Db 55 CAGGACATGCCAAGCCCTCAGAGCAGCGGTGCTCTCTGGCTTGGCCAGGCGACA 114
QY 70 CTGACCATCGCTTGGCTGTCTATCTTGGGCTCATGACCTTGGGCTTACGCGCGGG 129
Db 115 CTCAGATCATCTTTGGCTGTCTATCTCTGGGCTCATGACCTTGGGCTTACGCGCAGGC 174
QY 130 GTGCGCTGGCTCGCATGCTTACGCGCTCTCTGGCTTGGCTTACGCGCTTCTT 189

Db 175 GTTCCTCTGGCTTCAGATCGCTATGGAATCCTGGCTTTTGGCTCTATATGGGCAATTCCTC 234
QY 190 TCAGGGCACTGTGTGGCGCAGAGCCTCTTTCGGTACTTGGAGCAACCGCGGGTGGCGGCG 249
Db 235 AGCGCACACACTAGTGGCACAGAGCCTCTTTCGTTTACCTTGGAGCACCGAAGGGTGGCAGCG 294
QY 250 GCGGCGGCG-----GGGGCGCTGATGACGACCGCGCGCAGTGTGGCG 294
Db 295 GCTCGGCGGCGCTCTTGGCGAAGGGGCGCTGATGCGGCCACTGACGACGAGGTGGCA 354
QY 295 CTGACCATCTCCGCTTACCAGGAGGAGCCCGCGTACCTGCGCCAGTGTCTTGGCGTCCGCC 354
Db 355 CTACCATCTCAGCTTACCAGAGGATCCCGCTTACCTGCGCCAGTGTCTTGAACCTCCGCG 414
QY 355 CGCGCCCTGTGTATCCCGCGCGCGCTGCGCGCTCTCTCATGTGTGGTGGATGGCAACGCG 414
Db 415 CGCGCTTGTGTATCCCGCGCACACGAGGTTACGCGTGTCTCATGTGTGGTGGAGCGCAACGCG 474
QY 415 GCGGAGGACCTCTACATGTGTGACATGTTCCGCGAGGTCTTTCGCTGACGAGGACCCCGCG 474
Db 475 GCTGAGGATCTGTATGATGTGTGACATGTTCCGAGAGTCTTTCGCGGATGAGGACCCCGCG 534
QY 475 ACCTACGTGTGGGACGCGCACTACACAGCCCTTGGGAACCCCGCGCGG-----CGGGCGCG 531
Db 535 ACTTATGTGGGATGGCAACTACATCAGCCCTTGGGAACCCAGCGAGGCTTACGGGCGCT 594
QY 532 GTGGCGCGGAGGCTTATCGGAGGTGGAGCGGAGGATCTCTGGCGGCTTGGCAGTGGAG 591
Db 595 GTCGGTGAAGTGTCTTACCGGAGGTGGAGCGGAGGACCCCGCGCGGTTTGGCGGTGGAG 654
QY 592 GCGCTGTGTGAGGACTCGCAGGTGTGTGTGGCGCAGCGCTGSGGCGGCAAGCGCGAG 651
Db 655 GCGCTGTGTGAGAACACGAGAGTGTGTGTGGCTCAGCGCTTGGGCGGCAAGCGTGGAG 714
QY 652 GTCATGTACACAGCCTTCAAGCGCTCGGAGATTCTGGTGGACTACGTCGACGCTCTGTGAC 711
Db 715 GTCATGTACACAGCTTTCNAGGCACTGGGCGACTCCGTTGGACTAGCTGACGCTCTGTGAC 774
QY 712 TCGGACACAAGTTTGGACCCCATCGGCACTGTGTGAGCTCTGTGTGGGTTACTGGACGAGAC 771
Db 775 TCAGACACAAGACTAGACCCCATGGCACTGTGTGAGGCTTGTGCGAGTGTGTGATGAAGAC 834
QY 772 CCGCGGTAGGGCTGTGTGGGAGCTGTGGGATCTTAACTTAACTTAACTTAACTTAACTTAACT 831
Db 835 CCGCGGTAGGGCTGTGTGGGAGGATGTGAGGATCTTAACTTAACTTAACTTAACTTAACTTAACT 894
QY 832 AGCTTCTTAAAGCAGCTGCGATCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGAGAGC 891
Db 895 AGCTTCTTAAAGCAGCTTTCGATCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGAGAGC 954
QY 892 TACTTCCACTGTGTATCTTGCATCAGCGGTCTCTAGGCTTATATAGGAATAACCTCTTG 951
Db 955 TACTTCCACTGTGTATCTTGCATCAGCGGTCTCTAGGCTTATATAGGAATAACCTCTTG 1014
QY 952 CAGCAGTTCTTGGAGGCTGTGATCAACACAGAAATCTCTGGGTACCCACTGTACTTTTGGG 1011
Db 1015 CAGCAGTTCTTGGAGGCTGTGATCAACACAAAGTTCTCTGGGACCCACTGTGACATTTGGG 1074
QY 1012 GATGACCGGCACCTCACCACCGCATGCTCAGCATGGGTATGTACCAAGTA 1064
Db 1075 GATGACCGGCACCTCACCACCGCATGCTTAGCATGGGTATGTACCAAGTA 1127

RESULT 12
LOCUS MUSHAS

DEFINITION Mus musculus mRNA for hyaluronan synthase, complete cds.
ACCESSION D82964
VERSION D82964.1 GI:1339939
KEYWORDS hyaluronan synthase.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Matches 913; Conservative 0; Mismatches 1; Indels 141; Gaps 2;		
QY	10	CAGGACGGCCCAAGCCCACTCTCGAGCCCGCCGCTGCTCGGCTCGCCCGCGAGGGTG 69
DB	1	CAGGACGGCCCAAGCCCACTCTCGAGCCCGCCGCTGCTCGGCTCGCCCGCGAGGGTG 60
QY	70	CTGACCATCGCTTGGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCGCGCGG 129
DB	61	CTGACCATCGCTTGGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTACGCGCGCGG 120
QY	130	GTGCGCTGGCTCGCATCGCTACGCGCTCTGGGCTTGGGCTTCTACGGGCTTCTCTT 189
DB	121	GTGCGCTGGCTCGCATCGCTACGCGCTCTGGGCTTGGGCTTCTACGGGCTTCTCTT 180
QY	190	TCAGCGACCTGGTGCGAGAGCTCTTCGGGTACCTGGAGACCGGGGGGTGGCGGG 249
DB	181	TCAGCGACCTGGTGCGAGAGCTCTTCGGGTACCTGGAGACCGGGGGGTGGCGGG 212
QY	250	GCGCGCGGGGCGCTGGATGACGACCGCGCGAGTGTGGCGCTGACCATCTCCGCC 309
DB	213	-----TGTGGCGCTGACCATCTCCGCC 234
QY	310	TACGAGGAGACCCGCTACCTGGCCAGTGCCTGGGCTCGCCCGCGCTGCTGTAC 369
DB	235	TACGAGGAGACCCGCTACCTGGCCAGTGCCTGGGCTCGCCCGCGCTGCTGTAC 294
QY	370	CGCGCGCGCTGGCGCTCTCATGTGTGGATGGCAACCGCGCGGAGACCTCTAC 429
DB	295	CGCGCGCGCTGGCGCTCTCATGTGTGGATGGCAACCGCGCGGAGACCTCTAC 354
QY	430	ATGTGTGACATGCTTCGGGAGTCTTCGCTGACAGGAGCCCGCCACGTCGTGTGGAC 489
DB	355	ATGTGTGACATGCTTCGGGAGTCTTCGCTGACAGGAGCCCGCCACGTCGTGTGGAC 405
QY	490	GGCAACTACCAACAGCCCTGGGAACCCCGCGCGCGCGGCTGGCGCGGAGCCTAT 549
DB	406	----- 405
QY	550	CGGAGGTGGAGCGGAGATCCTGGCGGCTGGCAGTGGAGGCGCTGTTGAGGACTGCG 609
DB	406	-----GTGGAGCGGAGATCCTGGCGGCTGGCAGTGGAGGCGCTGTTGAGGACTGCG 459
QY	610	AGTGTGTGTGTGTCGCGCAGCGCTGGGCGCGCAAGCGCGAGTCTGTACACGCTTC 669
DB	460	AGTGTGTGTGTGTCGCGCAGCGCTGGGCGCGCAAGCGCGAGTCTGTACACGCTTC 519
QY	670	AAGCGGCTCGGAGATTCTGGTGACTACGTGCAAGTCTGTGACTCGGACACAGGTTGGAC 729
DB	520	AAGCGGCTCGGAGATTCTGGTGACTACGTGCAAGTCTGTGACTCGGACACAGGTTGGAC 579
QY	730	CCCATGGCACTGCTGGAGCTCTGGCGGTACTGGACGAGACCCCGCGGTAGGGCTGTT 789
DB	580	CCCATGGCACTGCTGGAGCTCTGGCGGTACTGGACGAGACCCCGCGGTAGGGCTGTT 639
QY	790	GSTGGGAGGTGGGATCTTAACCTCTGAGCTCTGGTCTGAGTCTTAAGCAGCTG 849
DB	640	GSTGGGAGGTGGGATCTTAACCTCTGAGCTCTGGTCTGAGTCTTAAGCAGCTG 699
QY	850	CGATACTGGGTAGCCTTCAATGTGGAGGGGCTTGTGAGAGTACTTCCACTGTGTATCC 909
DB	700	CGATACTGGGTAGCCTTCAATGTGGAGGGGCTTGTGAGAGTACTTCCACTGTGTATCC 759
QY	910	TGCATCAGGGTCTCTAGGCTTATAGGAATAACCTCTTGCAGAGTCTTCTTGAGGCC 969
DB	760	TGCATCAGGGTCTCTAGGCTTATAGGAATAACCTCTTGCAGAGTCTTCTTGAGGCC 819
QY	970	TGGTACACCAAGTCTCTGGTACCACTGTACTTTTGGGATGACCGGCACTCTACC 1029
DB	820	TGGTACACCAAGTCTCTGGTACCACTGTACTTTTGGGATGACCGGCACTCTACC 879
QY	1030	AACCGCATGCTCAGCATGGGTTATGCTACCAAGTA 1064
DB	880	AACCGCATGCTCAGCATGGGTTATGCTACCAAGTA 914

RESULT 14
AC018755/c
LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

REFERENCE

AUTHORS

JOURNAL

COMMENT

FEATURES

source

repeat_region

gene

mRNA

CDS

AC018755 229155 bp DNA linear PRI 26-JUL-2000
Homo sapiens chromosome 19, BAC BC330783 (CIT-HSPC_470E3), complete
sequence.

AC018755
AC018755.3 GI:9454515
HTG.

Homo sapiens (human)
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 229155)
Lamerdin,J.E., McCreedy,P.M., Skowronski,E., Viswanathan,V.,
Burkhardt-Schultz,K., Gordon,L., Dias,J., Ramirez,M., Stiliwagen,S.,
Phan,H., Velasco,N., Do,L., Regala,W., Terry,A., Brower,A.,
Garnes,J., Danganan,L., Erler,A., Christensen,M., Georgescu,A.,
Avila,J., Liu,S., Andreise,T., Frankheim,M., Attix,C.,
Amico-Keller,G., Coefield,J., Duarte,S., Lucas,S., Bruce,R.,
Thomas,P., Quan,G., Kronmiller,B., Arellano,A., Sanders,C., Ow,D.,
Nolan,M., Trong,S., Kobayashi,A., Olsen,A.S. and Carrano,A.V.
Sequence analysis of a 5-Mb region in 19q13.4

Unpublished
2 (bases 1 to 229155)
DOE Joint Genome Institute.
Direct Submission

Submitted (19-DEC-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 229155)
Lamerdin,J.E.

Direct Submission
Submitted (26-JUL-2000) DOE Joint Genome Institute, Lawrence
Livermore National Laboratory, Livermore, CA 94550
On Jul 26, 2000 this sequence version replaced gi:7458780.
Map and sequence oriented from centromere to q-telomere. BC330783
(CIT-HSPC_470E3) is overlapped on the left by BC849408
(CITB-EL_3073N11, AC020914) from bases 1 to 125 of this accession,
and overlaps comid R28782 (LLNL-R_248F10, AC005946) on the right
from bases 217,905 to 229,155 of this accession. Additional
chromosome 19 map and sequence information may be obtained at:
http://www.bio.llnl.gov/bbrp/genome/genome.html.

Location/Qualifiers
1. 229155
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
/clone="CIT-470E3 (LLNL clone name BC330783)"
/clone_lib="CalTech human BAC library C"
201. 262
/rpt_family="L2"
<299..16870
/gene="ZNF175"
/note="zinc finger protein 175; synonym: OTK18"
join(<299..550,8540..8666,8936..9031,13776..16870)
/gene="ZNF175"
/product="ZNF175 mRNA"
join(479..550,8540..8666,8936..9031,13776..15616)
/gene="ZNF175"
/codon_start=1
/product="zinc finger protein 175"
/protein_id="AAF87843.1"
/db_xref="GI:9454517"

/translation="MPADVNLSPKQVLPGEKDGSCASVSFDVTVDFSRBEWQOL
DPAGCLYRDMLELYSHLFAVGHIENPEVIFMLKEKPRVEAEVSHQRCORREF
GLETPKEIKKASQKQDMVGEFTRDGSGWGSILELRLDADRDKDEQNOIQPHSHA
FFNKKLNTESNCEYKDPKMMIRTPHLSAQKQPKCCLPFSLKLNLEVQNGNESN
DTEQLDDVSGQLFSSSDACSNHTGTFCQGNQKVKCVGHKQSLKQHDHTQK
KPDGSCGSGFTQKSHLFAQRIHSVGNLHECGKCGKAFMPQLKLSYLLDHTGDIIP
CICKECGKVFQSRSELLTHQKTRKPKYKCHDCGKAFQMLSLFRHQRTSHREKLYE


```
Db 146930 GGGGTGCGCTGGCGCTCCGATCGCTACGGCGCTCTCGGCTTTCGGCTCTACGGCGGCGCTTC 146871
QY 187 CTTTCAGCGCACCTGGTGGCGCAGAGCTCTTCGGTACCTGGAGCACCGCGGGTGGCG 246
Db 146870 CTTTCAGCGCACCTGGTGGCGCAGAGCTCTTCGGTACCTGGAGCACCGCGGGTGGCG 146811
QY 247 GCGGCGGCGCGGGGCGCGCTGGATGCACACCGCGCAGCTGTGGCGCTGACCATCTCC 306
Db 146810 GCGGCGGCGCGGGGCGCGCTGGATGCACACCGCGCAGCTGTGGCGCTGACCATCTCC 146751
QY 307 GCCTACCAAGGAGGACCGCGCTACCTGCGCAGTGCCTGGCGTCCGCGCGCGCGCTGCTG 366
Db 146750 GCCTACCAAGGAGGACCGCGCTACCTGCGCAGTGCCTGGCGTCCGCGCGCGCGCTGCTG 146691
QY 367 TACCGCGCGCGCGCTGGCTCTCATGTGTGGATGCAACCGCGCGCGGAGACTC 426
Db 146690 TACCGCGCGCGCGCTGGCTCTCATGTGTGGATGCAACCGCGCGCGGAGACTC 146631
QY 427 TACATGGTTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCGCGCACGTCGTGTGG 486
Db 146630 TACATGGTTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCGCGCACGTCGTGTGG 146571
QY 487 GACGCGCACTACCAACGAGCTTCGCTGACGAGGACCGCGCGCGCGCTGGCGCGCGGAGCC 546
Db 146570 GACGCGCACTACCAACGAGCTTCGCTGACGAGGACCGCGCGCGCGCTGGCGCGCGGAGCC 146511
QY 547 TATCGGAGGTGGAGCGGAGATCTTGGCGGCTGGCAGTGGAGGCGCTGGTGGAGACT 606
Db 146510 TATCGGAGGTGGAGCGGAGATCTTGGCGGCTGGCAGTGGAGGCGCTGGTGGAGACT 146451
QY 607 CGCAGGTGCGTGTGGTGGCGCAGCGCTGGCGCGCGCGCGCGCGCTGATGTACACAGCC 666
Db 146450 CGCAGGTGCGTGTGGTGGCGCAGCGCTGGCGCGCGCGCGCGCGCTGATGTACACAGCC 146391
QY 667 TTCAGGCGCTCGGAGATTCGGTGGACTACGTGACAGTCTGT 708
Db 146390 TTCAGGCGCTCGGAGATTCGGTGGACTACGTGACAGTCTGT 146349

RESULT 15
AC130783/c
LOCUS
DEFINITION Pan troglodytes clone CH251-426A12, WORKING DRAFT SEQUENCE, 8 linear HTG 19-NOV-2002
ordered pieces.
ACCESSION AC130783
VERSION AC130783.2 GI:25100968
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 193986)
Akhtar,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Cariaca,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
Laric,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.I., Maduro,V.B.,
Margulies,E.H., Masiello,C., Maskeri,B., McDowell,J.,
Paquirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
Reddix-Duque,N., Schandler,K., Schueler,M.G., Sison,C.,
Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
NISC Comparative Sequencing Initiative
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 193986)
Green,E.D.
JOURNAL Direct Submission
TITLE Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
REFERENCE 3 (bases 1 to 193986)
Green,E.D.
JOURNAL Direct Submission
TITLE Submitted (19-NOV-2002) NIH Intramural Sequencing Center, 8717
Grovemont Circle, Gaithersburg, MD 20877, USA
```

COMMENT

Grovemont Circle, Gaithersburg, MD 20877, USA
On Nov 19, 2002 this sequence version replaced gi:22218453.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: dxd
Center clone name: 426A12

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 191764 bases at least Q40
Consensus quality: 192510 bases at least Q30
Consensus quality: 192998 bases at least Q20
Insert size: 184000; agarose-fp
Quality coverage: 11.49x in Q20 bases; agarose-fp
Quality coverage: 10.94x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.

1 11512: contig of 11512 bp in length
* 11513 11612: gap of unknown length
* 11613 59488: contig of 47876 bp in length
* 59489 59588: gap of unknown length
* 59589 103266: contig of 43678 bp in length
* 103267 103366: gap of unknown length
* 103367 127885: contig of 24519 bp in length
* 127886 127985: gap of unknown length
* 127986 129692: contig of 1707 bp in length
* 129693 129792: gap of unknown length
* 129793 131077: contig of 1285 bp in length
* 131078 131177: gap of unknown length
* 131178 174663: contig of 43486 bp in length
* 174664 174763: gap of unknown length
* 174764 193986: contig of 19223 bp in length.

FEATURES

source

1. 193986
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="CH251-426A12"
/clone_lib="CH251"

misc_feature

1. 11512
/note="assembly_fragment
clone_end:SP6
vector_side:left"

misc_feature

1. .76
/note="clone overlaps with GenBank Accession Number
AC132069 clone CH251-132A6 (center project name dxc)"

```
misc_feature 11613..59488
/note="assembly_fragment"
misc_feature 59589..103266
/note="assembly_fragment"
misc_feature 103367..127885
/note="assembly_fragment"
misc_feature 127986..129692
/note="assembly_fragment"
misc_feature 129793..131077
/note="assembly_fragment"
misc_feature 131178..174663
/note="assembly_fragment"
misc_feature 174764..193986
/note="assembly_fragment
clone_end:T7
vector_side:right"

ORIGIN
Query Match 62.8%; Score 669; DB 2; Length 193986;
Best Local Similarity 97.2%; Pred. No. 9.2e-83;
Matches 697; Conservative 0; Mismatches 5; Indels 15; Gaps 1;

Qy 7 CAGCAGGAGCGCCCAAGCCCACTCTCGAGCCGCGCTGCTCGGCTGCGCCGCGAGG 66
Db 132828 CAGCAGGAGCGCCCAAGCCCACTCTCGAGCCGCGCTGCTCGGCTGCGCCGCGAGG 132769

Qy 67 GTGCTGACCATCGCTTCGCCCTTGCTCATCTGGGCTCATGACCTGGGCTTACGCCGCC 126
Db 132768 GTGCTGACCATCGCTTCGCCCTTGCTCATCTGGGCTCATGACCTGGGCTTACGCCGCC 132709

Qy 127 GGGGTGCGCTGCGCTTCGGATCGTACCGGCTCTCGGCTTCGGGCTTACCGGGCCTTC 186
Db 132708 GGGGTGCGCTGCGCTTCGGATCGTACCGGCTCTCGGCTTCGGGCTTACCGGGCCTTC 132649

Qy 187 CTTTCAGCGCACTGGTGGCGAGAGCTCTTCGGGTACCTGGAGACCGGCGGGTGGCG 246
Db 132648 CTTTCAGCGCACTGGTGGCGAGAGCTCTTCGGGTACCTGGAGACCGGCGGGTGGCG 132589

Qy 247 -----GCGCGCGCGCGCGCGCGCTGGATGACGACCGCGCGCAGTGTG 291
Db 132588 GCGCGCGCGCGCGCGCGCGCGCGCGCGCTGGATGACGACCGCGCGCAGTGTG 132529

Qy 292 GCGCTGACCATCTCGGCTTACAGAGGACCCCGGTACCTCGCGCAGTGCCTGGCGTCC 351
Db 132528 GCGCTGACCATCTCGGCTTACAGAGGACCCCGGTACCTCGCGCAGTGCCTGGCGTCC 132469

Qy 352 GCCCGCGCGCTGCTGTACCGCGCGCGCGCTGCGGCTCCTCATGGTGGTGGATGGCAAC 411
Db 132468 GCCCGCGCGCTGCTGTACCGCGCGCGCGCTGCGGCTCCTCATGGTGGTGGATGGCAAC 132409

Qy 412 CGCGCGGAGGACCTCTACATGGTGCACATGTTCCGCGAGTCTTCGCTGACGAGGACCCC 471
Db 132408 CGCGCGGAGGACCTCTACATGGTGCACATGTTCCGCGAGTCTTCGCTGACGAGGACCCC 132349

Qy 472 GCCACGTACGTGGGAGCGCAACTACACAGCCCTTGGGAACCCCGCGCGCGCGCGCG 531
Db 132348 GCCACGTACGTGGGAGCGCAACTACACAGCCCTTGGGAACCCCGCGCGCGCGCGCG 132289

Qy 532 GTGGCGCGCGGAGCTATCGGAGGTGGAGGCGGAGGATCTGGGCGGCTGGCAGTGGAG 591
Db 132288 GTGGCGCGCGGAGCTATCGGAGGTGGAGGCGGAGGATCTGGGCGGCTGGCAGTGGAG 132229

Qy 592 GGGCTGGTGGAGCTCGCAGGTGCGTGTGCGGCGAGCGCTGGGGCGGCAAGCGCGAG 651
Db 132228 GGGCTGGTGGAGCTCGCAGGTGCGTGTGCGGCGAGCGCTGGGGCGGCAAGCGCGAG 132169

Qy 652 GTCATGTACACAGCCTTCAAGCGCTCGAGATTCGGTGGACTACGTGCAGGTCTGT 708
Db 132168 GTCATGTACACAGCCTTCAAGCGCTCGAGATTCGGTGGACTACGTGCAGGTGAGT 132112
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 01:47:21 ; Search time 692.135 Seconds
(without alignments)
9108.809 Million cell updates/sec

Title: US-10-672-399-7

Perfect score: 1065

Sequence: 1 atgagacagcaggacgccc.....tgggttatgtaccacagtaa 1065

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1064	99.9	1737	ADL13691	Adl13691 Osteoarth
2	1064	99.9	2088	ADL13690	Adl13690 Osteoarth
3	1064	99.9	2116	AAT99541	Aat99541 Human hya
4	1062.4	99.8	1737	ADL13692	Adl13692 Osteoarth
5	1062.4	99.8	2087	ADL13694	Adl13694 Osteoarth
6	1042	97.8	2117	AAT96713	Aat96713 Human hya
7	786.6	73.9	1752	ABZ76734	Abz76734 Mouse hya
8	786.6	73.9	1752	AAD59442	Aad59442 Mouse hya
9	786.6	73.9	2102	AAT91655	Aat91655 Mouse hya
10	786.6	73.9	2102	AAT10862	Aat10862 Hyaluroa
11	786.6	73.9	2102	AAZ88199	Aaz88199 Mouse hya
12	786.6	73.9	2102	AAZ39987	Aaz39987 Murine HA
c 13	697.2	65.5	231222	ADL13693	Adl13693 Osteoarth
14	422	39.6	662	ABQ27384	Abq27384 Oligonuc
c 15	422	39.6	662	ABQ27385	Abq27385 Oligonuc
16	421	39.5	490	ACH39998	Ach39998 Human toe
c 17	393.8	37.0	662	ABQ27382	Abq27382 Oligonuc
18	393.8	37.0	662	ABQ27383	Abq27383 Oligonuc
19	366	34.4	1783	ADC49212	Adc49212 Rabbit hy
20	353	33.1	1662	ACC51023	Acc51023 Human bla

21	353	33.1	1662	8	ABX76415	Abx76415 Lung canc
22	353	33.1	1662	11	ADN39051	Adn39051 Cancer/an
23	353	33.1	4049	8	ABX76214	Abx76214 Lung canc
24	353	33.1	4049	12	ADN05115	Adn05115 Antipeori
25	353	33.1	4220	12	ADJ74905	Adj74905 Marker ge
26	323.8	30.4	1767	10	ADD93927	Add93927 Xenopus l
27	311.6	29.3	1665	2	AAV18822	Aav18822 Murine hy
28	311.6	29.3	1665	3	AZ88201	Aaz88201 Mouse hya
29	311.6	29.3	1665	8	ABZ76736	Abz76736 Mouse hya
30	311.6	29.3	1665	10	AAD59444	Aad59444 Mouse hya
31	311.6	29.3	5919	12	ADJ75814	Adj75814 Marker ge
c 32	246.8	23.2	1653	6	ABL41013	Ab141013 Murine hy
33	246.8	23.2	1659	8	ABZ76735	Abz76735 Mouse hya
34	246.8	23.2	1659	10	AAD59443	Aad59443 Mouse hya
35	246.8	23.2	2947	2	AAV18821	Aav18821 Murine hy
36	246.8	23.2	2948	6	ABI99697	Abi99697 Mouse isc
37	246.8	23.2	4194	9	ADA49685	Ada49685 DNA encod
38	246.8	23.2	4194	10	ADC59298	Adc59298 Mouse HAS
39	233.2	21.9	2890	3	AZ88200	Aaz88200 Mouse hya
40	221	20.8	1659	10	ADL13696	Adl13696 Osteoarth
41	221	20.8	3003	6	ABK84373	Abk84373 Human CDN
42	221	20.8	3003	9	ADA49683	Ada49683 DNA encod
43	221	20.8	3003	10	ADC59296	Adc59296 Human HAS
44	221	20.8	3003	10	ADD12718	Add12718 Human CDN
45	221	20.8	3003	10	ADL13695	Adl13695 Osteoarth

ALIGNMENTS

RESULT 1

ADL13691
ID ADL13691 standard; DNA; 1737 BP.

AC ADL13691;

DT 06-MAY-2004 (first entry)

DE Osteoarthritis-associated polymorphic nucleotide #223.

KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.

OS Homo sapiens.

FN WO20003054166-A2.

XX 03-JUL-2003.

XX 19-DEC-2002; 2002WO-US041225.

XX 20-DEC-2001; 2001US-0342603P.

XX (INCY-) INCYTE GENOMICS INC.

XX Jones KA, Schafer A;

XX WPI; 2003-559141/52.

XX Determining susceptibility of an individual to joint space narrowing,
XX osteophyte development and/or joint pain comprises identifying whether
XX the individual has at least one polymorphism in a polynucleotide encoding
XX a protein.

XX Disclosure; SEQ ID NO 223; 297pp; English.

XX The invention relates to a method of determining susceptibility of an
XX individual to joint space narrowing and/or osteophyte development and/or
XX joint pain comprising identifying whether the individual has at least one
XX polymorphism in a polynucleotide encoding at least one of the protein
XX listed in the specification. The methods, composition and agent are
XX useful for modulating the susceptibility of an individual to joint space

CC narrowing and/or osteophyte development and/or joint pain that is
CC associated with a disease, preferably osteoarthritis. The cell line and
CC the non-human animal are useful for screening for an agent for diagnosing
CC an individual having susceptibility to joint space narrowing and/or
CC osteophyte development and/or joint pain. This sequence corresponds to
CC the polynucleotide encoding a protein listed in the specification. (Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences).

XX SQ Sequence 1737 BP; 232 A; 571 C; 603 G; 331 T; 0 U; 0 Other;

Query Match 99.9%; Score 1064; DB 10; Length 1737;

Best Local Similarity 100.0%; Pred. No. 6.7e-180; Indels 0; Gaps 0;

Mismatches 0; Conservative 0; Matches 1064;

QY	1	ATGAGACAGCAGGAGCGCCCAAGCCCACTCTCTGAGCGCGCGCTGCTCGGCGCTGGCC	60
DB	1	ATGAGACAGCAGGAGCGCGCCCAAGCCCACTCTCTGAGCGCGCGCTGCTCGGCGCTGGCC	60
QY	61	CGGAGGGTGTGACCATCGCTTGGCTCTCTGATCTCTGAGCGCTCATGACCTGGGCGCTAC	120
DB	61	CGGAGGGTGTGACCATCGCTTGGCTCTCTGATCTCTGAGCGCTCATGACCTGGGCGCTAC	120
QY	121	GCCTCTCTTTCAGCGCACCTGCTGGCGCAGAGCCTCTTCGCGTACTCTGAGCAGCCGCGG	240
DB	181	GCCTCTCTTTCAGCGCACCTGCTGGCGCAGAGCCTCTTCGCGTACTCTGAGCAGCCGCGG	240
QY	241	GTGGCGGCGGCGCGCGCGCGCGCTGTGATGAGCAGCAGCGCGCAGTGTGGCGCTGAC	300
DB	241	GTGGCGGCGGCGCGCGCGCGCGCTGTGATGAGCAGCAGCGCGCAGTGTGGCGCTGAC	300
QY	301	ATCTCGGCTTACAGAGGAGCCCGCGTACCTGCGCCAGTGCCTGGCGCTGGCGCGCC	360
DB	301	ATCTCGGCTTACAGAGGAGCCCGCGTACCTGCGCCAGTGCCTGGCGCTGGCGCGCC	360
QY	361	CTGCTGTATACCCCGCGCGCGCTGCGCTCTCTCATGTGTGGATGGCAACCGCGCGG	420
DB	361	CTGCTGTATACCCCGCGCGCGCTGCGCTCTCTCATGTGTGGATGGCAACCGCGCGG	420
QY	421	GACCTCTATAGTGTGACATGTTCCGAGGTTCTTCGCTGACGAGGACCCCGCGCAGTAC	480
DB	421	GACCTCTATAGTGTGACATGTTCCGAGGTTCTTCGCTGACGAGGACCCCGCGCAGTAC	480
QY	481	GTGTGGGAGCGCACTACCAAGCCCTGGGAAACCGCGCGCGCGCGCGCTGGGCGCC	540
DB	481	GTGTGGGAGCGCACTACCAAGCCCTGGGAAACCGCGCGCGCGCGCGCTGGGCGCC	540
QY	541	GGAGCTTATCGGAGGTGGAGCGGAGGATCTCTGGCGCGCTGGCAGTGGAGCGCTGGTG	600
DB	541	GGAGCTTATCGGAGGTGGAGCGGAGGATCTCTGGCGCGCTGGCAGTGGAGCGCTGGTG	600
QY	601	AGGACTCGAGGT	660
DB	601	AGGACTCGAGGT	660
QY	661	ACAGCTTTCAGGCGCTCGGAGTTTCGGTGGACTACGTGCGAGTCTGTGACTCGGACACA	720
DB	661	ACAGCTTTCAGGCGCTCGGAGTTTCGGTGGACTACGTGCGAGTCTGTGACTCGGACACA	720
QY	721	AGTTTGGACCCCATGGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	780
DB	721	AGTTTGGACCCCATGGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	780
QY	781	GGGGT	840
DB	781	GGGGT	840
QY	841	AGCAGCTGCGATCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGTCAGAGCTACTTCCAC	900

DB	841	AGCAGCTGCGATCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGTCAGAGCTACTTCCAC	900
QY	901	TGTGTATCTGTCATCAGCGGTCCTCTAGGCTATATAGGAATAACCTCTTGAGCAGTTT	960
DB	901	TGTGTATCTGTCATCAGCGGTCCTCTAGGCTATATAGGAATAACCTCTTGAGCAGTTT	960
QY	961	CTTGAGGCTGTGTCACCAACCAAGTTCTCTGGGTACCCACTGTACTTTTGGGGATGACCGG	1020
DB	961	CTTGAGGCTGTGTCACCAACCAAGTTCTCTGGGTACCCACTGTACTTTTGGGGATGACCGG	1020
QY	1021	CACCTCACCAACCGCATGCTCAGCATGGGTATGTACCAAGTA	1064
DB	1021	CACCTCACCAACCGCATGCTCAGCATGGGTATGTACCAAGTA	1064

RESULT 2

ADL13690

ID ADL13690 standard; DNA; 2088 BP.

XX AC ADL13690;

XX DT 06-MAY-2004 (first entry)

XX OS Osteoarthritis-associated polymorphic nucleotide #222.

XX KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;

XX KW joint space narrowing; osteophyte development; joint pain;

XX KW osteoarthritis; SNP; single nucleotide polymorphism.

XX OS Homo sapiens.

XX PN WO2003054166-A2.

XX PD 03-JUL-2003.

XX PF 19-DEC-2002; 2002WO-US041225.

XX PR 20-DEC-2001; 2001US-0342603P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Jones KA, Schafer A;

XX DR WPI; 2003-559141/52.

XX CC Determining susceptibility of an individual to joint space narrowing, osteophyte development and/or joint pain comprises identifying whether the individual has at least one polymorphism in a polynucleotide encoding a protein.

XX PS Disclosure; SEQ ID NO 222; 297pp; English.

XX CC The invention relates to a method of determining susceptibility of an individual to joint space narrowing and/or osteophyte development and/or joint pain comprising identifying whether the individual has at least one polymorphism in a polynucleotide encoding at least one of the protein listed in the specification. The methods, composition and agent are useful for modulating the susceptibility of an individual to joint space narrowing and/or osteophyte development and/or joint pain that is associated with a disease, preferably osteoarthritis. The cell line and the non-human animal are useful for screening for an agent for diagnosing an individual having susceptibility to joint space narrowing and/or osteophyte development and/or joint pain. This sequence corresponds to the polynucleotide encoding a protein listed in the specification. (Note: The sequence data for this patent did not form part of the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences).

XX SQ Sequence 2088 BP; 302 A; 652 C; 711 G; 423 T; 0 U; 0 Other;

Query Match

Best Local Similarity 99.9%; Score 1064; DB 10; Length 2088;

Mismatches 0; Conservative 0; Matches 100.0%; Pred. No. 6.7e-180;

RESULT 3	
AAT99541	
ID	AAT99541 standard; cDNA; 2116 BP.
XX	
AC	AAT99541;
XX	
DT	21-MAY-1998 (first entry)
XX	
DE	Human hyaluronan synthase cDNA clone 30C.
XX	
KW	Hyaluronan synthase; HAS gene; human; hyaluronic acid; cell adhesion;
KW	wound healing; vulnerary; tissue repair; scar; keloid; therapy; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	CDS
FT	36..1772
FT	/*tag= a
FT	polyA_signal
FT	2066..2071
FT	/*tag= b
XX	
XX	
EN	W09740174-A1.
XX	
PD	30-OCT-1997.
XX	
PD	17-APR-1997; 97WO-US006350.
PF	
XX	
PR	22-APR-1996; 96US-00635552.
XX	
PA	(LEUK-) LEUKOSITE INC.
XX	
PI	Briskin MJ;
XX	
DR	WPI; 1997-549359/50.
DR	P-PSDB; AAW26765.
XX	
PT	Human hyaluronan synthase - useful for recombinant production of
PT	hyaluronic acid for wound healing, tissue repair and reducing
PT	hypertrophic scar and keloid formation.
XX	
PS	Claim 3; Page 36-38; 58pp; English.
XX	
CC	cDNA clone 30C includes a coding region for hyaluronan synthase (HAS)
CC	(see AAW26765) an enzyme involved in the synthesis of hyaluronan
CC	(hyaluronic acid) and which has the ability to confer cell adhesion by
CC	the lymphocyte receptor CD44. Clone 30C was isolated using an expression
CC	cloning system developed to isolate cDNA clones that encode proteins that
CC	confer adhesion of the murine T cell lymphoma TK1. A human mesenteric
CC	lymph node expression library was constructed that, upon transfection
CC	into CHO/P cells, yielded clone 30C that mediated reosetting of TK1 cells
CC	to some of the transfectants. The isolated clone can be utilised in a
CC	claimed method for producing HAS in host cells. Such host cells are used
CC	in a claimed method for the production of hyaluronan. Hyaluronan is
CC	useful for wound healing and tissue repair, and can reduce or prevent
CC	hypertrophic scars and keloid formation. It is also used in eye surgery
CC	as a replacement for vitreous fluid
XX	
SQ	Sequence 2116 BP; 330 A; 651 C; 712 G; 423 T; 0 U; 0 Other;
	Query Match 99.9%; Score 1064; DB 2; Length 2116;
	Best Local Similarity 100.0%; Pred. No. 6.7e-180;
	Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 ATGAGACAGCAGACGCGCCCAAGCCCACTCTCTGACGCCCGCGCTCTCGGCGCTGCC 60
Db	36 ATGAGACAGCAGACGCGCCCAAGCCCACTCTCTGACGCCCGCGCTCTCGGCGCTGCC 95
Qy	61 CGAGGGTGCTGACCATCGCCTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
Db	96 CGAGGGTGCTGACCATCGCCTTCGCCCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 155

Db 301 ATCTCCGCTTACAGAGGAGACCCCGCTACCTGCGCCAGTGCCTGGCGCTCCGCGCGCC 360
Qy 361 CTGCTGTATACCGCGCGCGCTGGCGCTCTCATGTGGTGGATGGCAACCGCGCCGAG 420
Db 361 CTGCTGTATACCGCGCGCGCTGGCGCTCTCATGTGGTGGATGGCAACCGCGCCGAG 420
Qy 421 GACCTCTACATGGTGCAGATGTCGCGAGGTCCTCGCTGACGAGACCCCGCCACGTAC 480
Db 421 GACCTCTACATGGTGCAGATGTCGCGAGGTCCTCGCTGACGAGACCCCGCCACGTAC 480
Qy 481 GTGTGGAGCGCACTACACAGCCCTGGGAAACCGCGCGCGCGCGCGCTGGCGCC 540
Db 481 GTGTGGAGCGCACTACACAGCCCTGGGAAACCGCGCGCGCGCGCGCTGGCGCC 540
Qy 541 GGAGCTATCGGAGGTGGAGCGGAGATCTGCGCGGCTGGCAGTGGAGCGCTGGTG 600
Db 541 GGAGCTATCGGAGGTGGAGCGGAGATCTGCGCGGCTGGCAGTGGAGCGCTGGTG 600
Qy 601 AGGACTCGAGGTGCTGTGGTGGCGAGCGCTGGCGCGCAACGCGAGGTCATGTAC 660
Db 601 AGGACTCGAGGTGCTGTGGTGGCGAGCGCTGGCGCGCAACGCGAGGTCATGTAC 660
Qy 661 ACAGCTTCAAGCCCTCGGAGATTCGCTGAGTACGTCAGGTCGTGTGACTCGACACA 720
Db 661 ACAGCTTCAAGCCCTCGGAGATTCGCTGAGTACGTCAGGTCGTGTGACTCGACACA 720
Qy 721 AGTTGGAGCCCATGCGACTGCTGAGCTCGTGGGCTACTGGAGAGGACCCCGGGTA 780
Db 721 AGTTGGAGCCCATGCGACTGCTGAGCTCGTGGGCTACTGGAGAGGACCCCGGGTA 780
Qy 781 GGGGCTGTGGTGGGAGCGTGGGATCCTTAACCTCTGGAATCTCTGGGTGAGTTCCTA 840
Db 781 GGGGCTGTGGTGGGAGCGTGGGATCCTTAACCTCTGGAATCTCTGGGTGAGTTCCTA 840
Qy 841 AGCAGCTGCATACCTGAGTACCTCAATGTGAGCGGCTTGTGAGACTACTTCAC 900
Db 841 AGCAGCTGCATACCTGAGTACCTCAATGTGAGCGGCTTGTGAGACTACTTCAC 900
Qy 901 TGTGTATCTGATCAGCGGCTCTAGGCTTATAGGAATAAATCTTTCAGAGATT 960
Db 901 TGTGTATCTGATCAGCGGCTCTAGGCTTATAGGAATAAATCTTTCAGAGATT 960
Qy 961 CTGAGGCTGGTACAAACAGAGTTCTGGGTACCCACTGTACTTTTGGGATGACCGG 1020
Db 961 CTGAGGCTGGTACAAACAGAGTTCTGGGTACCCACTGTACTTTTGGGATGACCGG 1020
Qy 1021 CACCTCACAACCGCATGCTGAGCATGGGTTATGCTACCAAGTA 1064
Db 1021 CACCTCACAACCGCATGCTGAGCATGGGTTATGCTACCAAGTA 1064

RESULT 5
ADL13694

ID ADL13694 standard; DNA; 2087 BP.

XX AC ADL13694;

XX DT 06-MAY-2004 (first entry)

XX DE Osteoarthritis-associated polymorphic nucleotide #226.

XX KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
joint space narrowing; osteophyte development; joint pain;
XX KW osteoarthritis; SNP; single nucleotide polymorphism.

XX OS Homo sapiens.

XX PN WO2003054166-A2.

XX PD 03-JUL-2003.

XX PF 19-DEC-2002; 2002WO-US041225.

XX

PR 20-DEC-2001; 2001US-0342603P.

XX (INCY-) INCYTE GENOMICS INC.

PI Jones KA, Schafer A;

XX WPI; 2003-559141/52.

Determining susceptibility of an individual to joint space narrowing,
osteophyte development and/or joint pain comprises identifying whether
the individual has at least one polymorphism in a polynucleotide encoding
a protein.

Disclosure; SEQ ID NO 226; 297pp; English.

The invention relates to a method of determining susceptibility of an
individual to joint space narrowing and/or osteophyte development and/or
joint pain comprising identifying whether the individual has at least one
polymorphism in a polynucleotide encoding at least one of the protein
listed in the specification. The methods, composition and agent are
useful for modulating the susceptibility of an individual to joint space
narrowing and/or osteophyte development and/or joint pain that is
associated with a disease, preferably osteoarthritis. The cell line and
the non-human animal are useful for screening for an agent for diagnosing
an individual having susceptibility to joint space narrowing and/or
osteophyte development and/or joint pain. This sequence corresponds to
the polynucleotide encoding a protein listed in the specification. (Note:
The sequence data for this patent did not form part of the printed
specification but was obtained in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences).

SQ Sequence 2087 BP; 302 A; 650 C; 712 G; 423 T; 0 U; 0 Other;

Query Match 99.8%; Score 1062.4; DB 10; Length 2087;

Best Local Similarity 99.9%; Pred. No. 1.3e-179; Indels 0; Gaps 0;
Matches 1063; Conservative 0; Mismatches 1;

Qy 1 ATGAGACAGCAGGAGCGGCCCAAGCCCACTCTCTGACGCCCGCGCTGCTCGGCGCTGCGCC 60

Db 36 ATGAGACAGCAGGAGCGGCCCAAGCCCACTCTCTGACGCCCGCGCTGCTCGGCGCTGCGCC 95

Qy 61 CGAGGGTGTGACATGCGCTTCCGCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120

Db 96 CGAGGGTGTGACATGCGCTTCCGCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 155

Qy 121 GCCCGCGGGTGGCGCTGGCGCTCCGATCGCTACCGGCTCTCGGCTTCTACCGG 180

Db 156 GCCCGCGGGTGGCGCTGGCGCTCCGATCGCTACCGGCTCTCGGCTTCTACCGG 215

Qy 181 GCCTTCTCTTTCAGCGCACCTGGTGGCGCAGAGCCTCTTTCGGCTTCTGGAGCACCGGCGG 240

Db 216 GCCTTCTCTTTCAGCGCACCTGGTGGCGCAGAGCCTCTTTCGGCTTCTGGAGCACCGGCGG 275

Qy 241 GTGCGCGCGCGCGCGCGCGCGCGCTGATGACGCCACCGCGCGAGTGTGGCGCTGACC 300

Db 276 GTGCGCGCGCGCGCGCGCGCGCGCTGATGACGCCACCGCGCGAGTGTGGCGCTGACC 335

Qy 301 ATCTCCGCTTACAGGAGGACCCCGCTTACCTGGCGCAGTGGCGTGGCGCTCGCGCGCC 360

Db 336 ATCTCCGCTTACAGGAGGACCCCGCTTACCTGGCGCAGTGGCGTGGCGTGGCGCTCGCGCC 395

Qy 361 CTGCTGTATACCGCGCGCGCGCTGCGCTCTCATGTGGTGGATGGCAACCGCGCCGAG 420

Db 396 CTGCTGTATACCGCGCGCGCGCTGCGCTCTCATGTGGTGGATGGCAACCGCGCCGAG 455

Qy 421 GACCTCTACATGGTGCAGATGTCGCGAGGTCCTGCGCTGACGAGGACCCCGCCACGTAC 480

Db 456 GACCTCTACATGGTGCAGATGTCGCGAGGTCCTGCGCTGACGAGGACCCCGCCACGTAC 515

Qy 481 GTGTGGAGCGCACTACCAACCGCCCTGGGAAACCGCGCGCGCGCGCGTGGCGCC 540

Db 516 GTGTGGAGCGCACTACCAACCGCCCTGGGAAACCGCGCGCGCGCGCGTGGCGCC 575

QY 772 CCCGGTAGGGCTGTGGTGGGACCTGGGATCCTTAACCTCTGGACTCCCTGGGTC 831
 Db 787 CCCGGTAGGGCTGTGGAGGGATGTGAGGATCTTAACCTCTGGACTCCCTGGGTC 846
 QY 832 AGCTTCCCTAAGCAGCTCGATACCTAGGCTTCAATGTGGAGCGGGCTTGTGAGAGC 891
 Db 847 AGCTTCTTGAGCAGCTCTGATACCTAGGCTTCAATGTGGAGCGGGCTTGTGAGAGC 906
 QY 892 TACTTCCACTGTGTATCTTGCATACGCGGTCTCTAGGCTTATATAGGAATAACCTCTTG 951
 Db 907 TACTTCCACTGTGTCTGCTGATCAGTGGTCTCTGGGCTTATATAGGAATAACCTCTTG 966
 QY 952 CAGCAGTTCTTGGAGGCTGTGACACAGAGTTCTGGGTACCCATGTACTTTGGG 1011
 Db 967 CAGCAGTTCTTGGAGGCTGTGACACCAAGTTCTTGGGACCCACTGACATTTGGG 1026
 QY 1012 GATGACCGGCACCTCAACACCGCATGCTCAGCATGGGTATGTACCAAGTA 1064
 Db 1027 GATGACCGGCACCTCAACACCGAATGTCTAGCATGGGCTATGCTACCAAGTA 1079

RESULT 8

ID AAD59442 standard; DNA; 1752 BP.

AC AAD59442;

DT 18-DEC-2003 (first entry)

DE Mouse hyaluronan synthase (HAS) 1 DNA.

KW Mouse; hyaluronic acid; HA; hyaluronan synthase; HAS; dry eye syndrome;

OS Mus sp.

Key Location/Qualifiers

1..1752

/tag= a

/product= "Mouse hyaluronan synthase 1"

US2003087850-A1.

08-MAY-2003.

10-JUL-2001; 2001US-00902939.

10-JUL-2001; 2001US-00902939.

(DEHA/) DEHAZYA P.

(CHEN/) CHEN W.

Dehazy P, Chen W;

WPI; 2003-755151/71.

P-P5DB; AAE39152.

Dihydrate derivatized hyaluronic acid/nucleic acid bioconjugate

comprising derivatized hyaluronic acid cross linked to nucleic acid

encoding hyaluronan synthase useful for treating dry eye syndrome.

Claim 1; Page 12-13; 31pp; English.

The invention relates to dihydrate derivatized hyaluronic acid (HA) /

nucleic acid bioconjugate comprising derivatized HA cross linked to

nucleic acid encoding hyaluronan synthase (HAS). The invention is useful

for transfecting an eye cell of an individual. It is useful for treating

dry eye syndrome and osteoarthritis of the particular joints. The

invention is also useful in gene therapy. The present sequence is mouse

HAS1 DNA

Sequence 1752 BP; 308 A; 505 C; 548 G; 391 T; 0 U; 0 Other;

Query Match 73.9%; Score 786.6; DB 10; Length 1752;
 Best Local Similarity 84.9%; Pred. No. 1.2e-130; Indels 18; Gaps 2;
 Matches 911; Conservative 0; Mismatches 144;
 QY 10 CAGGACGGCCCAAGACCCACCTCCCTGCAGCCCGCGCTGCTCCGGCTTGGCCCGGAGGGTG 69
 Db 7 CAGGACATGCCAAGACCCCTCAGAGGAGCGGCTGCTCTGCGCTTGGCCCGGAGGAGCA 66
 QY 70 CTGACCATTCGCTTTGGCCCTGCTCATCTGCGGCTCTATGACCTTGGGCTTACGCGCCGGG 129
 Db 67 CTACGATCATCTTTGGCCCTGCTCATCTGCGGCTCTATGACCTTGGGCTTACGCGCAGGC 126
 QY 130 GTGCGCTGGGCTCCGATCGCTAGCGGCTTCTGCGCTTCCGGCTTACGCGGCTTCTTCTT 189
 Db 127 GTTCTCTTGGCTTTCAGATCGCTATGGACTTCTGCGCTTTCGGCTTTCATGCGGCTTCTC 186
 QY 190 TCAGCGCACCTGTGTGGGCGCAGAGCTCTTTCCGCTACCTTGGAGCACCGCGGCTTGGCGCG 249
 Db 187 AGCGCACACCTAGTGGCACAGAGCTCTTTCGCTTACCTTGGAGCACCGGAGGCTGGCAGG 246
 QY 250 GCGGCGG-----GGGCGCTGATGACGACCCGCGCGCAGTGTGGCG 294
 Db 247 GCTGCGGCGCTCTTTGGCGAAGGGGCCCTTGTATGCGGCCACTGTCACGACGCTGGCA 306
 QY 295 CTGACCATCTCCGCTTACCGAGGAGGACCCCGCGTACCTTGGCCAGTCTGCGCTCGCGCC 354
 Db 307 CTCACCATCTCAGCTTACCGAGGATCCCGCTTACCTTGGCCAGTGTGACCTCCGCG 366
 QY 355 CGCGCTCTGTGTACCGCGCGCGCTGCGCTCTCTCATGCTGTGGTGGATGGCAACCGC 414
 Db 367 CGCGCTTGTGTACCGCGCACACGAGTTACGCGTCTCATGCTGTGGTGGACGCAACCGC 426
 QY 415 GCCAGGACCTCTACATGCTGCAATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCC 474
 Db 427 GCTGAGGATCTGTATGCTGTGCAATGTTCCGAGAGTCTTTCGCGCATGAGGACCCCGCC 486
 QY 475 ACCTACCTGTGGAGCGGCACTACACAGCCCTGGGNAACCGCGCGG-----CGGCGCG 531
 Db 487 ACTTATGTGTGGAGTGGCACTACATCAGCCCTTGGGNAACCGCGGAGGCTACGCGCGCT 546
 QY 532 GTGGCGCGCGAGCTTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTTGGCAGTGGAG 591
 Db 547 GTCGTGAAGGTGCTTACCGGAGGTGGAGCGGAGGACCCCGCGGCTTGGCGTGGAG 606
 QY 592 GCGCTGTGAGGACTCGCAGGTGCTGTGCGCGCGAGCGCTTGGGCGGCGCAAGCGCAG 651
 Db 607 GCGCTGTGAGGACACGCGAGGTGCTGTGCGTGGCTCAGCGCTTGGGCGCGCAACGCTGAG 666
 QY 652 GTCATGTACACAGCTTCAAGCGCTCGGAGATTCGGTGGACTAGTCAGCTCTGTGAC 711
 Db 667 GTCATGTACACAGCTTCAAGCGACTGGCGGACTCCCGTGGACTAGTCAGAGTCTGTGAC 726
 QY 712 TCGGACACAGGTGGGACCCCATGGCACTGTGGAGCTCTGTGGGCTTACGACGAGGAC 771
 Db 727 TCAGACACAGACTAGACCCCATGCGACTGTGGAGCTTGTGCGAGTGTGGATGAAGAC 786
 QY 772 CCCCGGTAGGGCTGTGTGGGAGCGTGGGATTCCTTAACTCTTGGACTCTCGGCTC 831
 Db 787 CCCCGGTAGGGCTGTGTGGAGGGGATGTGAGGATTCCTTAACTCTTGGACTCTCGGCTC 846
 QY 832 AGCTTCTTAGCAGCTTCCGATACCTGCGTACCTTCAATGTGGAGCGGCTTGTTCAGAGC 891
 Db 847 AGCTTCTTAGCAGCTTCCGATACCTGCGTACCTTCAATGTGGAGCGGCTTGTTCAGAGC 906
 QY 892 TACTTCCACTGTGTATCTTGCATCAGCGGCTCTCTAGGCTTATATAGGAATAACCTCTTG 951
 Db 907 TACTTCCACTGTGTCTTGCATCAGTGGTCTCTGCGTCTTATACAGAAACAATCTCCTG 966
 QY 952 CAGCAGTTCTTGGAGGCTGTGTAACAACAGAGTTCTTGGGTACCCACTGTACTTTGGG 1011
 Db 967 CAGCAGTTCTTGGAGGCTGTGTAACAACAAAGTTCTTGGGCGACCCCACTGCACATTTGGG 1026

Qy 1012 GATGACCGGCACCTCACCACCGCATGTCTCAGCATGGGTTATGCTACCAAGTA 1064
 |||||
Dδ 1027 GATGACAGGCACCTCACCAACCGAATGCTTAGCATGGGGTATGCTACCAAGTA 1079

RESULT 9

AAT91655
ID AAT91655 standard; DNA; 2102 BP.

AA91655;

	17-OCT-2003	(revised)
DT	19-DEC-1997	(first entry)
DT		

DE Mouse hyaluronate synthase genomic DNA.

KW Hyaluronate synthase; hyaluronic acid; polysaccharide; ds.

OS Mus sp; (strain FM3A):

FT	Key	Location/Qualifiers
FH	CDS	49. .1800
FT		/*tag= a
FT		/product= "Hyaluron

JP09224674-A.

02-SEP-1997.

26-FEB-1996: 96JP-00038336.

PR 26-FEB-1996; 96JP-00038336.

PA (KAGG) KAGAKU GIJITSU CHO CHOKAN KANBO.

DR WPI; 1997-484102/45.

DR P-PSDB; AAW30704.

PT Hyaluronate synthase isolated from mouse cells - useful for large-scale production of hyaluronic acid.

PS Claim 2; Page 10-13; 15pp; Japanese.

A mouse-derived cell strain, FM3A P15-A, was cultured in Eagle minimal medium containing 10 % heat-inactivated bovine serum, twice concentration of amino acids and vitamins and penicillin and streptomycin at 37 degrees Celsius. The culture was subjected to immobilised erythrocyte exclusion to examine the extent of extracellular formation of hyaluronic acid matrix. Cells which showed high formation were recovered and named FM3A HAl. Cells synthesising no hyaluronic T antigen were recovered and named HAS-. The HAS- cell in which polyoma large T antigen was expressed was prepared (HAS- P cell). Poly(A)+ RNA was isolated from FM3A HAl and cDNA was prepared and was used for constructing a library in HAS- P cells. Cells having hyaluronic acid synthetic activity were selected from the transformants and plasmid DNA was recovered and amplified in E.coli. The resulting genomic DNA sequence codes for hyaluronate synthase having a sequence of 583 amino acids which is used for large-scale production of hyaluronic acid, a component of drugs and cosmetics. (Updated on 17-OCT-2003 to standardise OS field)

Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;

Query Match 73.9%; Score 786.6; DB 2; Length 2102;

Best Local Similarity 84.9%; Pred. No. 1.2e-130;
Matches 911; Conservative 0; Mismatches 144; Indels 18; Gaps 2;

QY 10 CAGGACGGCGCCCAAGCCCACTCTCTGCAGCCCGCGCTGTCTCGGCTGGCCCGAGGGTG 69

Db 55 CAGGACATGCCAAAGCCCTCAGAGGCAGCGCTTGCTCTGGCCTGGCCAGCGAGCA 114

QY 70 CTGACCATCGCCCTTCGCCCTGCTCATCCTGGGCTCATGACCTGGGCTACGCCCGGG 129

D_b 115 CTCACGATCATCTTTGCCCTGCTCATCTGGGCTCATGACCTGGCTACGCCGAGGC 174

QY	130	GTGCGCTGGCGTCCGATCGGTACGGCCTCTCGGCCCTTACGGGGCTTCCTT	189
Db	175	GTTCCTCTGGCTTCAGATCGTATGGACTCTCGCCCTTTGGCCCTCTATGGGCAATTCCTC	234
QY	190	TCAGCGACCTGTGTGGCGCAGAGCTCTTCGCGTACTGTGAGCACCAGCGGGTGGCGGCG	249
Db	235	AGCGCACACCTTAGTGGCACAGAGCCTCTTCGCTTACTGTGAGACCAAGAGGTGGACGC	294
QY	250	GGCGCGCG-----GGGGCCGCTGGATGACGCCACCGCGCGCAGTGTGGCG	294
Db	295	GCTGCGCGCGCTCTTGTGGCAAGGGGCGCCCTGGATGCGGCCACTGCA CGCAGCGTGSCA	354
QY	295	CTGACCAATCTCGGCTACAGAGGAGACCCGCGGTACTGTGCGCAGTGCCTTGGGGTTCGCGC	354
Db	355	CTCACCAATCTCAGCTTACCAAGAGGATCCCGCTTACTCTGCGCCAGTGTCTTGACCTCGCG	414
QY	355	CGCGCCCTGTGTACCCGCGCGCGCGCTCGCGTCTCATGTGTGGTGGATGGCAACCGC	414
Db	415	CGCGCCCTGTGTACCCGCGCACACGAGGTTAGCGGTGCTCATGTGTGTGGACGGCAACCGC	474
QY	415	GCGGAGGACCTCTACATGCTCGATGTTCCGCGAGGTCTTCGTGACGAGGACCCCGCC	474
Db	475	GCTGAGGATCTGTACATGGTGGACATGTTCCGAGAAAGTCTTCCGCGATGAGGACCCCGCC	534
QY	475	AGTACTGTGTGGAGCGCAACTACACAGCCCTCGGAAACCCGCGCGCG---CGGGCGCG	531
Db	535	ACTTATGTGTGGATGGCAACTACCATCAGCCCTGGGAAACAGCGGAGGCTACGGCGCT	594
QY	532	GTGGGCGCGGAGCTATCGGGAGGTGGAGCGGAGGATCTCGGGCGCTGGGAGTGGAG	591
Db	595	GTGCGTGAAGTGTCTACCGGAGGTGGAGCGGAGGACCCCGCGCGGTTGGCGGTGGAG	654
QY	592	GCGCTGTGTAGGACTCGCAGGTGCGTGTGCGTGGCGCAGCGCTGGGGCGCAAGCGCGAG	651
Db	655	GCGCTGTGTAGAAACAGCGAGTGTGCTGTGGCTCAGCGTTGGGCGCGCAACGTTGAG	714
QY	652	GTCAATGTACACAGGCTTCAAGGCGCTCGGAGATTCCGTGGAATCTCGTCAAGTCTGTGAC	711
Db	715	GTCAATGTACACAGCTTTCAAGGCACCTGGGCGACTCCGTGGACTACGTGCAAGTCTGTGAC	774
QY	712	TCGGACACAAAGTTGGACCCCATGGCACTCTCGAGCTCGTGGCGGTACTGGACAGGAGAC	771
Db	775	TCAGACACAAGACTAGACCCCATGGCACTGCTGGAGCTTGTGCGAGTTGTGGATGAAGAC	834
QY	772	CCCGGGTAGGGCTGTGTGTGGGGACGTGCGGATCCTTAAACCTCTGGACTCCTGGGTC	831
Db	835	CCCGGGTAGGGCTGTGTGGAGGGGATGTGAGGATCCTTAAACCTCTGGACTCCTGGGTC	894
QY	832	AGCTTCTTAAGACAGCTTGGGATACTGGGTAGCCTTCAATGTGAGCGGGCTTGTCAAGC	891
Db	895	AGCTTCTTGAAGCAGTCTTCGATATGGGTAGCCTTCAATGTGGAAACGAGCTTGTCAAGC	954
QY	892	TACTTCCACTGTGTATCTCATCAGCGGTCCCTAGGCCCTATATAGGAATAACCTCTTG	951
Db	955	TACTTCCACTGTGTGTCTGATCAGTGTGCTCTGGGTCTATACAGAAACAATCTCTTG	1014
QY	952	CAGCAGTTTCTTGAGGCTGGTACAACCGAAGTCTCTGGGTACCCACTGTACTTTGGG	1011
Db	1015	CAGCAGTTCTTGAGGCGCTGTACAACCAAAAGTTCTTGGGCAACCCACTGCACATTTGG	1074
QY	1012	GATGACCGGCACCTCACCAACCGCGATGCTCAGCATGGGTTATGCTACCAAGTA	1064
Db	1075	GATGACAGGCACCTCACCAACCGAATGCTTAGCATGGGCTATGCTACCAAGTA	1127

RESULT 10
AAZ10862
ID . AAZ10862 standard; DNA; 2102 BP.

AA
AC AAZ10862;

DT 19-OCT-1999 (first entry)

10 CAGACGCGCCACAGCCATCTCTGCAGCCGCGCTCTCGGCGCTGGCCGAGGCTG 59
Db CAGACATCCAAAGCCCTCAGAGCAGCGGTTGCTCTCTGGCTCGCCAGGCGAGCA 114
70 CTGACCATCGCCTTCGCGCTCTCATCTCGGCGCTCATGACTGGGCTAAGCCGCCCGG 129
Db CTCACGATCATCTTTGGCTCTCATCTCGGCGCTCATGACTGGGCTAAGCCCGAGC 174

XX Hyaluronate synthase coding sequence.
DE
XX Hyaluronate synthase; promoter; mouse; cell growth inhibitor; ds.
KW
XX Mus sp.
OS
XX Key Location/Qualifiers
FT CDS 49..1800
FT /*tag= a
XX JP11196875-A.
XX 27-JUL-1999.
XX 14-JAN-1998; 98JP-00006191.
XX 14-JAN-1998; 98JP-00006191.
XX (SEK) SEIKAGAKU KOGYO CO LTD.
XX WPI; 1999-496653/42.
DR P-PSDB; AAY32503.
XX New promoter DNA of hyaluronate synthase - used to specifically express
PT gene with cell growth inhibiting activity.
XX
XX Example 1; Page 8-10; 13pp; Japanese.
XX This sequence represents the mouse hyaluronate synthase coding sequence,
CC and was used to isolate the hyaluronate synthase promoter of the
CC invention. The promoter can be used for specifically expressing a gene
CC having cell growth inhibiting activity
XX
SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;
Query Match 73.9%; Score 786.6; DB 2; Length 2102;
Best Local Similarity 84.9%; Pred. No. 1.2e-130;
Matches 911; Conservative 0; Mismatches 144; Indels 18; Gaps 2;
QY 10 CAGGACGGCCCAAGCCCACTCCTGCAGCCCGCCGCTGCTCGGCTCGCCGAGGGTG 69
DB 55 CAGGACATGCCAAGCCCTCAGAGGAGCGGCTGCTGCTCGGCTCGCCAGCGAGCA 114
QY 70 CTGACCATGCGCTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCTCATGCGCCGG 129
DB 115 CTCAGATCATCTTTGCCCTGCTCATCTGGGCTCATGACCTGGGCTCATGCGCCAGC 174
QY 130 GTGCGCTGGCTCGATCGCTACGCGCTCTCGGCTTCGGCTCTACGGGCTTCCTT 189
DB 175 GTTCTCTGGCTTCAGATCGCTATGACCTCTGGGCTTTGGGCTCTATGGGCAATTCCTC 234
QY 190 TCAGCGCACCTGGTGGCGCAGAGCTCTTTGGGTACCTGGAGCACCGGGGGTGGCGCG 249
DB 235 AGCGCACACCTAGTGCGACAGAGCTCTTCGTTACTTGGAGCACCGAAGGTGGCAGCG 294
QY 250 GGGGGCGG-----GGGGCGCTGGATGACGACCGCGCTGAGCGCTTCGCGCG 294
DB 295 GCTGCGCGCGCTCTCTGGCGAAGGGGGCCCTGGATGGGCGCACTGCACGCGGTGGCA 354
QY 295 CTGACCATCTCGCTTACCGAGGAGCACCGCGGTACCTGGCGCTGCGCTGGCGCGCC 354
DB 355 CTCACCATCTCAGCTTACCAAGAGATCCCGTTTACCTTGGCGGCTTGAACCTTCGCG 414
QY 355 CGCGCCCTGCTGTATCCGCGCGCGGCTGCGGCTCTCATGCTGGTGGATGGCAACCGC 414
DB 415 CGGCGCTTGTCTACCCGACACAGAGTTACGCGTCTCATGGTGTGGACGGCAACCGC 474
QY 415 GCGGAGGACCTTACATGGTGAATGTTTCGCGAGGCTTCGCTGACGAGGACCCCGCC 474
DB 475 GCTGAGGATCTGTATCATGTGGACATGTTCCGAGAAGTCTTCGCGGATGAGGACCCCGCC 534
QY 475 ACGTACGTGTGGAGCGGCACTTACCACGAGCCCTGGGAACCCGCGCGG---CGGGCGG 531

DB 535 ACTTATGTGGATGGCACTACCATCAGCCCTGGGAACCAAGCGGCTACGGGCGCT 594
QY 532 GTGGCGCCGAGGCTATCGGAGGTGGAGCGGAGGATCTTGCGCGCTGGCAGTGGAG 591
DB 595 GTCGTGAAGGTGCTTACCGGGAGGTGGAGCGGAGGACCCCGGCGGTTGGCGTGGAG 654
QY 592 GCGCTGGTGAAGACTCCGAGGTGCGTGTGCGTGGCGAGCGCTGGGGCGGAGCGCGAG 651
DB 655 GCGTGGTGAAGAACACGAGGTGCGTGTGCGTCTAGCGTTGGGGCGGCAACGTCGAG 714
QY 652 GTCATGTACACAGGCTTCAAGCGCTCGGAGATTCGGTGGACTACGTCAGGTCTGTGAC 711
DB 715 GTCATGTACACAGCTTTCAGGCACTGGGCGACTCCGTGGACTAGTCCAGGTCTGTGAC 774
QY 712 TCGGACACAAAGTTGGACCCCATCGCACTGTGTGAGCTGTGCGGGTACTGGACGAGGAC 771
DB 775 TCAGACACAAAGTAGACCCCATGCGACTGTGGAGCTTGTGCGAGTGTGGATGAAGAC 834
QY 772 CCGCGGTAGGGCTGTGGTGGGACGTGGGATCCTTAACCTCTGGACTCCTGGGTC 831
DB 835 CCGCGGTAGGGCTGTGGAGGGGATGTGAGGATCCTTAACCTCTGGACTCCTGGGTC 894
QY 832 AGCTTCTTAAGCAGCTGCGGATCTGGGTAGCCTTCAATGTGGAGCGGCTTGTGACAGC 891
DB 895 AGCTTCTTGACAGCTTTCGATCTGGGTAGCCTTCAATGTGGAGCGGCTTGTGACAGC 954
QY 892 TACTTCCACTGTGTATCCTGCATCAGCGGCTCTTAGGCTTATAGGAATAACCTCTTG 951
DB 955 TACTTCCACTGTGTGCTGCTGCATCAGTGTCTCTGGGCTTATACAGAAACAATCTCCTG 1014
QY 952 CAGCAGTTCTTGAGGCTGTGTACACAGAGTTCCTGGTACCCTGCTGCTGCTGCTGGG 1011
DB 1015 CAGCAGTTCTTGGAGGCTGTGTACACCAAAAGTTCCTGGGCGCCACTGTCACATTTGGG 1074
QY 1012 GATCAGCGGACCTCACCAACCGCATGCTCAGCATGGTATGCTTACCAGTA 1064
DB 1075 GATCAGCAGGACCTCACCAACCGCATGCTTACGATGGCTATGCTACCAAGTA 1127
RESULT 11
AAZ88199
ID AAZ88199 standard; cDNA to mRNA; 2102 BP.
XX AAZ88199;
DT 27-APR-2000 (first entry)
XX Mouse hyaluronate synthase modified protein HAS1 encoding cDNA.
XX Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; ds.
XX Mus sp.
XX Key Location/Qualifiers
FT CDS 49..1800
FT /*tag= a
FT /product= "HAS1"
FT /note= "modified hyaluronate synthase protein"
XX JP2000004886-A.
XX 11-JAN-2000.
XX 24-JUN-1998; 98JP-00193788.
XX 24-JUN-1998; 98JP-00193788.
XX (SEK) SEIKAGAKU KOGYO CO LTD.
XX WPI; 2000-140125/13.
DR P-PSDB; AAY68491.

XX A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.
PS
XX Claim 18; Page 12-14; 30pp; Japanese.

CC The present invention describes a recombinant protein which consists of
CC three continuous regions (N-terminal region, internal region and C-
CC terminal region) where one or two regions among the above three regions
CC is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. The present sequence encodes mouse HAS1

XX SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;

Query Match 73.9%; Score 786.6; DB 3; Length 2102;
Best Local Similarity 84.9%; Pred. No. 1.2e-130;
Matches 911; Conservative 0; Mismatches 144; Indels 18; Gaps 2;

QY 10 CAGGACGGCCCAAGCCCACTCTCTCAGCCGCGCTCTCGGCTCGSCCGGAGGGTG 69
DB 55 CAGGACATGCCAAGCCCTCAGAGGAGCGCGCTTCTGCTCTGGCTGGCCAGGCGAGCA 114
QY 70 CTGACCATCGCCTTCGCGCTCTCATCTCTGGGCTCATGACCTGGGCTTACGCCCGCGG 129
DB 115 CTCAGATCATCTTTGGCTCTCATCTCTGGGCTCATGACCTGGGCTTACGCCCGCAGGC 174
QY 130 GTGCGCTGGCTTCGATCGCTTACCGCTCTCTGGCTCTTGGCTCTTACGGGCTTCTT 189
DB 175 GTTCTCTGGCTTCAGATCGCTATGAGCTCTCTGGCTTGGCTCTTATGGGCTTCTTCTC 234
QY 190 TCAGCGCACCTGGTGGCGCAGAGCTCTTCTGGTACCTTGGGCTTGGGCTTGGGCTG 249
DB 235 AGCGCACACCTAGTGGCAGAGCTCTTCTGGTACCTTGGGCTTGGGCTTGGGCTG 294
QY 250 GCGGCGCG-----GCGGCGCTTGGATGACGACCGCGCTGAGTGTGGCG 294
DB 295 GCTGGCGCGCTCTTGGCGAAGGCGCTTGGATGCGGCTGACGACGAGCTGGCA 354
QY 295 CTGACCATCTCGCTTACAGAGAGACCGCGCTTACCTGCGCACTGCTGGCGTCCGCC 354
DB 355 CTCACCATCTCAGCTTACAGAGAGATCCCGCTTACCTGCGCACTGCTTACCTCCGCG 414
QY 355 CGCGCGCTGCTTACCGCGCGCGCTGCGGCTCTCATGCTGGTGGTGGTGGTGGTGGTGG 414
DB 415 CGCGCTTGTGTATCCCGCACAGAGGTTACGCGTGTCTCATGCTGGTGGTGGTGGTGGTGG 474
QY 415 GCCGAGGACCTCTACATGCTGACATGTTCCGCGAGGTCTTCTGCTGACGAGGACCGCGCC 474
DB 475 GCTGAGGATCTTACATGCTGACATGTTCCGAGAGTCTTCTGCGCGATGAGGACCGCGCC 534
QY 475 AGCTACGTTGGGACCGCAACTTACCAAGCGCTTGGGAAACCGCGCGCG-----CGGCGCG 531
DB 535 ACTTATGTTGGGATGGCAACTTACCATCAGCCCTGGGAACCAAGCGGAGCTACGCGCGCT 594
QY 532 GTGGCGCGGAGCTATCGGAGGTGGAGCGGAGGATCTTGGGCGCTTGGCAGTGGAG 591
DB 595 GTGCGTGAAGGTGCTTACCGGAGGTGGAGCGGAGGACCGCGCGCTTGGCGTGGAG 654
QY 592 GCGCTGTGTGAGGACCTCGAGGTGCTGTGCTGGGCGAGCGCTGGGCGGCAAGCGCGAG 651
DB 655 GCGCTGTGTGAGNACAGGAGGTGCTGTGCTGGTGTGAGCTTGGGCGGCAAGCTGAG 714
QY 652 GTCATGTACAGCGCTTCAAGCGCTCGAGATTCGGTGGTACCTGCTCAGGTCTGTGAC 711
DB 715 GTCATGTACAGCGCTTCAAGCGCTCGAGATTCGGTGGTACCTGCTCAGGTCTGTGAC 774
QY 712 TCGGNACAGAGTTGGACCCCATGGCACTGTGGAGCTGTGGGCTTGGTGGTGGTGGTGG 771
DB 775 TCAGACACAGACTAGACCCCATGGCACTGTGGAGCTTGTGGAGCTTGTGGAGTGTGGAGT 834

QY 772 CCCGGGTAGGGGCTTGTGTGGGACGTGCGGATCTTAAACCCCTCTGGACTCTGGGTG 831
DB 835 CCCGGGTAGGGGCTTGTGTGGGAGGATGTGAGGATCTTAAACCCCTCTGGACTCTGGGTG 894
QY 832 AGCTTCCTAAGCAGCGCTCGATACCTGGGTAGGCTTCAATGTGGAGCGGCTTGTGAGAGC 891
DB 895 AGCTTCCTTGGAGCAGTCTTTCGATACCTGGGTAGGCTTCAATGTGGAGCGGCTTGTGAGAGC 954
QY 892 TACTTCCACTGTGTATCTCGCATCAGCGGTCTCTTAGGCCCTATATAGGAATAACCTCTTG 951
DB 955 TACTTCCACTGTGTCTCTCGCATCAGGTCTCTCTGGGTCTATACAGAAACATCTCTTG 1014
QY 952 CAGCAGTTTCTTGGGCTCTGTACAAACAGAGTTCTCTGGGTACCCACTGTACTTTTGG 1011
DB 1015 CAGCAGTTTCTTGGGCTCTGTACAAACAGAGTTCTCTGGGTACCCACTGTGACATTTGGG 1074
QY 1012 GATGACCGGACCTCACCACCGCATGCTAGCATGGTATGTACCAAGTA 1064
DB 1075 GATGACAGGCACTCACCACCGCATGCTAGCATGGTATGTACCAAGTA 1127

RESULT 12

AAA39987
ID AAA39987 standard; cDNA; 2102 BP.

XX AC AAA39987;

XX DT 19-OCT-2000 (first entry)

XX DE Murine HAS1 cDNA.

XX KW Murine; HAS1; hyaluronate synthase; marker gene; intracellular loop;
XX gene targeting; ss.

XX OS Mus sp.

XX FH Key Location/Qualifiers
XX CDS 49..1800
XX FT /*tag= a
XX FT /product= "HAS1"

XX PN JP2000116382-A.

XX PD 25-APR-2000.

XX PF 13-OCT-1998; 98JP-00291201.

XX PR 13-OCT-1998; 98JP-00291201.

XX PA (SEBK) SEIKAGAKU KOGYO CO LTD.

XX WPI; 2000-369404/32.

XX P-PSDB; AAB09948.

XX DNA used for gene targeting of the hyaluronate synthase gene (HAS1).

XX Example; Page 7-9; 14pp; Japanese.

XX This invention describes a novel DNA which contains an heterologous DNA
CC (I) to be introduced to chromosomal DNA of a host cell, a first and
CC second homologous region DNA (II) connected respectively to the 5' and 3',
CC sides of the introduced DNA, and a negative marker gene (III) expressible
CC in the host cell. (I) contains a positive marker gene expressible in the
CC host cell. (I), (II) and the region encoding the intracellular loop of
CC the hyaluronate synthase (HAS) 1 protein from chromosomal DNA is
CC constituted so as to cause a homologous recombination. The DNA can be
CC used for gene targeting of the HAS1 gene. This sequence encodes the
CC murine HAS1 protein described in the method of the invention

XX SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;

Query Match 73.9%; Score 786.6; DB 3; Length 2102;

QY 187 CTTTCAGCGCACCTGGTGGCGAGAGCCTCTTCGGGTACTGTGAGCACCGCGGGTGGCG 246
Db 148942 CTTTCAGCGCACCTGGTGGCGAGAGCCTCTTCGGGTACTGTGAGCACCGCGGGTGGCG 148883
QY 247 GCGGCGCGCGGGCGCGTGTGATGACGACCGCGCGAGTGTGGCGTGAACATCTCC 306
Db 148882 GCGGCGCGCGGGCGCGTGTGATGACGACCGCGCGAGTGTGGCGTGAACATCTCC 148823
QY 307 GCTTACCAGGAGACCCCGGTACCTGCGCAGTGCCTGGCGTCCGCGCGCGCTGCTG 366
Db 148822 GCTTACCAGGAGACCCCGGTACCTGCGCAGTGCCTGGCGTCCGCGCGCGCTGCTG 148763
QY 367 TACC CGCGCGCGCTGCGGCTCTCATGTGGTGGATGGCAACCGCGCGCGAGACCTC 426
Db 148762 TACC CGCGCGCGCTGCGGCTCTCATGTGGTGGATGGCAACCGCGCGCGAGACCTC 148703
QY 427 TACATGTTGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCGCACGTGCTGG 486
Db 148702 TACATGTTGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCGCACGTGCTGG 148643
QY 487 GACGCACTACACAGCCCTGGGAACCCGCGCGCGCGCGTGGCGCGCGAGCC 546
Db 148642 GACGCACTACACAGCCCTGGGAACCCGCGCGCGCGCGTGGCGCGCGAGCC 148583
QY 547 TATCGGAGGTGGAGCGGAGGATCCTTGGCGGCTGGCAGTGGAGCGCTGGTGAAGACT 606
Db 148582 TATCGGAGGTGGAGCGGAGGATCCTTGGCGGCTGGCAGTGGAGCGCTGGTGAAGACT 148523
QY 607 GCAGGTGCGTGTGCGCGCAGCGCTGGGCGCGCAAGCGCGAGGTGATGACAGCC 666
Db 148522 GCAGGTGCGTGTGCGCGCAGCGCTGGGCGCGCAAGCGCGAGGTGATGACAGCC 148463
QY 667 TTCAGGCGCTCGGAGATTCGCTGACCTACGTCGAGTCTGT 708
Db 148462 TTCAGGCGCTCGGAGATTCGCTGACCTACGTCGAGTCTGT 148421

RESULT 14
ID ABO27384
XX ABO27384 standard; DNA; 662 BP.
XX AC ABO27384;
XX DT 12-JUL-2002 (first entry)
XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 13975.
XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
XX KW drug; side effect; cancer; central nervous system; cardiovascular;
XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;
XX KW SNP; cell differentiation; ds.
XX OS Homo sapiens.
XX PN WO200218632-A2.
XX PD 07-MAR-2002.
XX PF 01-SEP-2001; 2001WO-EP010074.
XX PR 01-SEP-2000; 2000DE-01043826.
XX PR 05-SEP-2000; 2000DE-01044543.
XX PA (EPIC-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX DR WPI; 2002-371829/40.
XX PT Determining the degree of cytosine methylation in genomic DNA, useful for
PT diagnosis and prognosis, comprises selective hybridization of amplicons
PT from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.
XX This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridized to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABO13410-
CC ABO54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
XX
SQ Sequence 662 BP; 80 A; 80 C; 247 G; 255 T; 0 U; 0 Other;
Query Match 39.6%; Score 422; DB 6; Length 662;
Best Local Similarity 77.3%; Pred. No. 6.4e-66;
Matches 512; Conservative 0; Mismatches 150; Indels 0; Gaps 0;
QY 34 GCAGCCCGCCCTGCTCCGGCTGCGCCCGGAGGGTGTGACCATCGCCTTCGCCCTGCTC 93
Db 1 GTAGTTTCTGCTGTTTTCGGTTTGGTTCGAGGGTGTGATTATCGTTTTCGTTTGTGTT 60
QY 94 ATCTGGGCTCATGACCTGGGCTACCGCGCGGGGTGCGCGCTCCGATCGCTAC 153
Db 61 ATTTTGGGTTTATGATTGGTGTACGTCGCGGGTGTGCTGTTGGTTTCGATCGTTTAC 120
QY 154 GGCTCTCTGGGCTTCGGCTCTAGCGGCTTCCTTTTCAGCGCACCTCGTGGCGAGAGC 213
Db 121 GGTTTTTCGGTTTTCGGTTTTCGGGTTTTCGGGTTTTCGGGTTTTCGGGTTTTCGGGTT 180
QY 214 CTCTTCGCTACCTGGAGCACCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 273
Db 181 TTTTTCGCGTATTTGGAGTATCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240
QY 274 GGCACCG 333
Db 241 GTTATCGCGCGTAGTGTGGCGTTGATTATTTTCGTTTATTAGGAGGATTCGCGTATTG 300
QY 334 CGCCAGTGCCTGGCGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTC 393
Db 301 GGTAGTGTTCGGGCTTCGTTTCGGGTTTTCGGGTTTTCGGGTTTTCGGGTTTTCGGGTTT 360
QY 394 ATGTTGTTGGATGGCAACCGCGCGAGGACCTCTACATGGTTCGATGTTTCGCGAGGTC 453
Db 361 ATGTTGTTGGATGGTAACTCGCGTTCGAGGATTTTATATGTCGATATGTTTCGCGAGGT 420
QY 454 TTCGCTGACAGGAGACCCCG 513
Db 421 TTCGTTGACGAGGATTTTCGTTTACGTTACGTTACGTTACGTTTATTTATTTAGTTTGG 480
QY 514 CCCGCTC 573
Db 481 TTCGCTT 540
QY 574 GGGCGCTTCGCGAGTGGAGCGCGCTGCTGAGGACTCGCAGGTGCTGTGCTGCGCGAGGCG 633
Db 541 GGGCGCTTCGTTAGTGGAGCGCGTTCGTTAGGATTCGTTAGTTCGTTGCTGCTGCGTAGCGT 600
QY 634 TGGGCGCGCAAGCGCGAGGTGATGTAACAGCCCTTCAAGCGCGCTCGAGAGATTCGGTGGAC 693
Db 601 TGGGCGCGTAAGCGCGAGGTTATGTATATAGTTTTCGAGGCGTTTCGAGAGATTCGGTGGAT 660

Query Match:
Best Local
Matches 10

Qy	70	CTGACCATCGCTTGGCCCTGCTCATCTCTGGCCCTCATGACCTGGGCTACGCCGCGGG	129
Ds	118	CTGACCATCGCTTGGCCCTGCTCATCTCTGGCCCTCATGACCTGGGCTACGCCGCGGG	177
Qy	130	GTGCGCTGGCTCGATCGCTACGGCTCTCTGGCTTCTGGCTTCTAGGGGCTTCTCTT	189
Ds	178	GTGCGCTGGCTCGATCGCTACGGCTCTCTGGCTTCTGGCTTCTAGGGGCTTCTCTT	237
Qy	190	TCAGCGCACTGTGGCGCAGAGCTCTTCGCTGCTACCTGGAGCAACCGCGGCTGGCGCG	249
Ds	238	TCAGCGCACTGTGGCGCAGAGCTCTTCGCTGCTACCTGGAGCAACCGCGGCTGGCGCG	297
Qy	250	GCGCGCGCGGCGCTGGATGACGACCGCGGCGAGTGTGGCTGACCATCTCGCC	309
Ds	298	GCGCGCGCGGCGCTGGATGACGACCGCGGCGAGTGTGGCTGACCATCTCGCC	357
Qy	310	TACAGAGGACCGCGCTGCTGCGCAGTGTCTGGCTCGCGCGCGCTCTGTGTAC	369
Ds	358	TACAGAGGACCGCGCTGCTGCGCAGTGTCTGGCTCGCGCGCGCTCTGTGTAC	417
Qy	370	CCGCGCGCGGCTGCGCTCTCATGTGTGTGATGCAACCGCGCGCGAGCACTCTAC	429
Ds	418	CCGCGCGCGGCTGCGCTCTCATGTGTGTGATGCAACCGCGCGCGAGCACTCTAC	477
Qy	430	ATGCTCGACATGTTCCGCGAGTCTTCTGCTGACGAGGACCGCGCACTAGTGTGGAC	489
Ds	478	ATGCTCGACATGTTCCGCGAGTCTTCTGCTGACGAGGACCGCGCACTAGTGTGGAC	537
Qy	490	GGCAACTACCAACGAGCTTGGGAAACCGCGCGCGCGCTGGGCGCGGAGCTTAT	549
Ds	538	GGCAACTACCAACGAGCTTGGGAAACCGCGCGCGCGCTGGGCGCGGAGCTTAT	597
Qy	550	CGGAGGTGAGGCGGAGGATCTTGGGCGGCTGCGAGTGGAGGCGCTGGTGAAGCACTGC	609
Ds	598	CGGAGGTGAGGCGGAGGATCTTGGGCGGCTGCGAGTGGAGGCGCTGGTGAAGCACTGC	657
Qy	610	AGGTGCTGTGCTGGCGCAGGCTTGGGCGGCGGCGGAGGCTCATGTACAGGCTTC	669
Ds	658	AGGTGCTGTGCTGGCGCAGGCTTGGGCGGCGGCGGAGGCTCATGTACAGGCTTC	717
Qy	670	AAGGCGCTCGAGATTCGGTGACTACTGTCAGGCTCTGACTCGGACACAAAGGTTGGAC	729
Ds	718	AAGGCGCTCGAGATTCGGTGACTACTGTCAGGCTCTGACTCGGACACAAAGGTTGGAC	777
Qy	730	CCCATGGCACTGCTGAGCTCTGCGGCTACTGACGAGGACCCCGGGTAGGGGCTGTT	789
Ds	778	CCCATGGCACTGCTGAGCTCTGCGGCTACTGACGAGGACCCCGGGTAGGGGCTGTT	837
Qy	790	GCTGGGACGTCGGATCTTAAACCTCTGACTCTGGGTCAGCTTCTTAAGCAGCTG	849
Ds	838	GCTGGGACGTCGGATCTTAAACCTCTGACTCTGGGTCAGCTTCTTAAGCAGCTG	897
Qy	850	CGATCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGGCTACTTCCACTGTGTATCC	909
Ds	898	CGATCTGGGTAGCTTCAATGTGGAGCGGCTTGTGAGGCTACTTCCACTGTGTATCC	957
Qy	910	TGCTACAGCGCTCTTAGGCTATATAGGAATAACCTTTGACGAGATTTCTTGAAGCC	969
Ds	958	TGCTACAGCGCTCTTAGGCTATATAGGAATAACCTTTGACGAGATTTCTTGAAGCC	1017
Qy	970	TGCTACAGCAAGTTCCTGGTACCCACTGTACTTTTGGGATGACCGGCACTTACC	1029
Ds	1018	TGCTACAGCAAGTTCCTGGTACCCACTGTACTTTTGGGATGACCGGCACTTACC	1077
Qy	1030	AACCGCATGCTCAGCATGGGTTATGTACCAAGTA 1064	
Ds	1078	AACCGCATGCTCAGCATGGGTTATGTACCAAGTA 1112	

DEFINITION	Homo sapiens HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.		
ACCESSION	AY421569		
VERSION	AY421569.1 GI:39748428		
KEYWORDS	GSS.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 1728)		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios		
JOURNAL	Science 302 (5652), 1960-1963 (2003)		
PUBMED	14671302		
REFERENCE	2 (bases 1 to 1728)		
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA		
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.		
FEATURES	Location/Qualifiers		
source	1..1728		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
gene	<1..>1728		
	/gene="HAS1"		
	/locus_tag="HCM7601"		
ORIGIN			
Query Match	98.8%;	Score 1051.8;	DB 9; Length 1728;
Best Local Similarity	99.8%;	Pred. No. 8.8e-210;	
Matches 1053;	Conservative 0;	Mismatches 2;	Indels 0; Gaps 0;
Qy	10	CAGGACGCGCCCAAGCCACTCTCTGACGCGCGCGCTGCTCCGCGCTGGCCCGGAGGGTG	69
Ds	1	CAGGACGCGCCCAAGCCACTCTCTGACGCGCGCGCTGCTCCGCGCTGGCCCGGAGGGTG	60
Qy	70	CTGACCATCGCTTTCGCCCTGCTCATCTCGGGCTCATGACCTGGGCTACGCCGCGGG	129
Ds	61	CTGACCATCGCTTTCGCCCTGCTCATCTCGGGCTCATGACCTGGGCTACGCCGCGGG	120
Qy	130	GTGCGCTGGCTCGATCGCTACGGCTCTCGGCTTTCGGCTTTCGGGCTTTCCTT	189
Ds	121	GTGCGCTGGCTCGATCGCTACGGCTCTCGGCTTTCGGCTTTCGGGCTTTCCTT	180
Qy	190	TCAGCGCACTGTGGCGCAGAGCTCTTCGCTGCTGAGACCGCGCGGCTGGCGCG	249
Ds	181	TCAGCGCACTGTGGCGCAGAGCTCTTCGCTGCTGAGACCGCGCGGCTGGCGCG	240
Qy	250	GCGCGCGGGGCGCTGGATGACGACCGCGCGAGTGTGGCGCTGACCATCTCGCC	309
Ds	241	GCGCGCGGGGCGCTGGATGACGACCGCGCGAGTGTGGCGCTGACCATCTCGCC	300
Qy	310	TACAGAGGACCGCGGCTACCTGCGCAGTGTGCGCTTTCGGCTTTCGGCTTTCGTAC	369
Ds	301	TACAGAGGACCGCGGCTACCTGCGCAGTGTGCGCTTTCGGCTTTCGGCTTTCGTAC	360
Qy	370	CCGCGCGCGGCTGCGGCTCTCATGTGTGTGATGGCAACCGCGCGGAGCACTCTAC	429
Ds	361	CCGCGCGCGGCTGCGGCTCTCATGTGTGTGATGGCAACCGCGCGGAGCACTCTAC	420
Qy	430	ATGCTCGACATGTTTCGCGAGGCTTTCGCTGACGAGGACCCCGCACTAGTGTGGAC	489
Ds	421	ATGCTCGACATGTTTCGCGAGGCTTTCGCTGACGAGGACCCCGCACTAGTGTGGAC	480


```
Db      835  GNGGGGAGCTGGGAACTTAAACCTCTGGACTCTGGTCACTTCTTAAAGCAACCTG 894
Qy      850  CGAT-CTGGGAGCTTCAAT-GTGGAGCGGCTTG 884
Db      895  CGAANAAGTGGGTAGCTTCAATGTTGGAACGGGCTTG 931

RESULT 4
AL545447      965 bp      mRNA      linear      EST 25-MAR-2004
LOCUS      AL545447 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION      clone CS0D1027YJ02 5-PRIME, mRNA sequence.
ACCESSION      AL545447
VERSION      AL545447.3 GI:45745929
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:31267282.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9502.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?&CS0D1027DE01QP1&c=9502.r.
FEATURES
source
1..965
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1027YJ02"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match      74.7%; Score 795.6; DB 1; Length 965;
Best Local Similarity 97.4%; Pred. No. 3.6e-156;
Matches 851; Conservative 9; Mismatches 8; Indels 6; Gaps 5;
Qy      10  CAGGACGGCCCAAGCCACTCTCGACGCGCGCGCTGCTCGGCTGGCCGCGGAGGTG 69
Db      58  CAGGACGGCCCAAGCCACTCTCGACGCGCGCGCTGCTCGGCTGGCCGCGGAGGTG 116
Qy      70  CTGACCATCGCTTCGCCCTGCTCATCTGCGCTCATGACCTGGGCTACGCGCGCGG 129
Db      117  CTGACCATCGCTTCGCCCTGCTCATCTGCGCTCATGACCTGGGCTACGCGCGCGG 176
Qy      130  GTGGCGCTGGCTTCGATCGCTACGGCTCTGGGCTTCGGCTCTACGGGCGCTTCCTT 189
Db      177  GTGGCGCTGGCTTCGATCGCTACGGCTCTGGGCTTCGGCTCTACGGGCGCTTCCTT 236
Qy      190  TCAGGCGAAGCTGTGGCGAGAGCTCTTCGGCTACCTGGAGCACCGCGGGGTGGCGCG 249
Db      237  TCAGGCGAAGCTGTGGCGAGAGCTCTTCGGCTACCTGGAGCACCGCGGGGTGGCGCG 296
Qy      250  GCGGCGGGGCGCGCTGGTGGAGCACCGCGCGAGTGTGGCGCTGACCATCTCCGCC 309
Db      297  GCGGCGGGGCGCGCTGGATGCAGCCACGCGCGCAGTGTGGCTGTGACCATCTCCGCC 356
Qy      310  TACAGAGAGACCCCGGCTACCTGGCCAGTGCCTTGGGCTCCGCGCGGCGCTGTGTAC 369
Db      357  TACAGAGAGACCCCGGCTACCTGGCCAGTGCCTTGGGCTCCGCGCGGCGCTGTGTAC 416
Qy      370  CCGGCGCGCGCTGCGCTCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 429
Db      417  CCGGCGCGCGCTGCGCTCTCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 476
Qy      430  ATGTGTGACATGTTTCCGCGAGGCTTTCGTCGACAGGAGACCCCGCACGTACGTGTGGAC 489
Db      477  ATGTGTGACATGTTTCCGCGAGGCTTTCGTCGACAGGAGACCCCGCACGTACGTGTGGAC 536
Qy      490  GGCAACTACCAAGCCCTTGGGAAACCGCGCGCGCGCGCGCGCGCGCGCGAGCCTAT 549
Db      537  GGCAACTACCAAGCCCTTGGGAAACCGCGCGCGCGCGCGCGCGCGCGCGAGCCTAT 596
Qy      550  CCGGAGGTGGAGGCGGAGGATCCCTGGCGCGCTGCGAGTGGAGGCGCTGGTGAGACTCGC 609
Db      597  CCGGAGGTGGAGGCGGAGGATCCCTGGCGCGCTGCGAGTGGAGGCGCTGGTGAGACTCGC 656
Qy      610  AGGTGCGTGTGCGTGGCGCGAGCGCTGGGCGCGCAAGCGCGAGGTCACTACACAGCCTTC 669
Db      657  AGGTGCGTGTGCGTGGCGCGAGCGCTGGGCGCGCAAGCGCGAGGTCACTACACAGCCTTC 716
Qy      670  AAGCGCTCGAGATTCGTTGGGACTACGTGAGGTCTGTGACTCGGACACAAAGTTGGAC 729
Db      717  AAGCGCTCGAGATTCGTTGGGACTACGTGAGGTCTGTGACTCGGACACAAAGTTGGAC 776
Qy      730  CCCATGCACTGCTGGAGCTGCTGCGGCTACTGACGAGGACCCCGGGTAGGGGCTGTT 789
Db      777  CCCATGCACTGCTGGAGCTGCTGCGGCTACTGACGAGGACCCCGGGTAGGGGCTGTT 836
Qy      790  GGTGGGACGTGGCGATCCTTAACCTCTGAGCTCTCGGCTGAGCTTCCTAAGCAGCCTG 849
Db      837  GGTGGGACGTGGCGAT-CYTWAMCCTTGGACTCTCTKGGTAR--CTTCTAAGCAG-CTG 892
Qy      850  CGATACCTGGGTAGCCTTCAATGTGGAGCGGCTT 883
Db      893  CGATACCTGGGTAG-CTTYAATGTGGAGCGGCTTK 925

RESULT 5
AY421571      1746 bp      DNA      linear      GSS 17-DEC-2003
LOCUS      Mus musculus HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION      genomic survey sequence.
ACCESSION      AY421571
VERSION      AY421571.1 GI:39748430
KEYWORDS      GSS.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1746)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 1746)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submision
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
COMMENT
```

```
them based on alignment.
FEATURES             Location/Qualifiers
  source              1..1746
                     /organism="Mus musculus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10090"
  gene                <1..>1746
                     /gene="HAS1"
                     /locus_tag="HCM7601"

ORIGIN
Query Match          74.0%; Score 788.2; DB 9; Length 1746;
Best Local Similarity 85.0%; Pred. No. 1.3e-154;
Matches 912; Conservative 0; Mismatches 143; Indels 18; Gaps 2;

QY 10 CAGGACGGCCCAAGCCACTCTCGAGCCCGCCGCTGCTCGGCTGCGCCGAGGGTG 69
    |||||
Db 1 CAGGACATGCCAAGCCCTCAGAGCGAGCGCTTGTGCTTGGCTTGGCCAGCGGCA 60

QY 70 CTGACCATGGCTTCGCCCTGCTCATCTCCGCTCATGACCTGGGCTTACGCGCCGG 129
    |||||
Db 61 CTCAGATCATCTTTGCCCTGCTCATCTCGGCTCATGACCTGGGCTTAGCGCGAG 120

QY 130 GTGCGCTGGCTCCGATCGCTACGGCTCTCTGGCTTGGCTCTAAGGGCTTCTT 189
    |||||
Db 121 GTTCTCTGGCTTCAGATCGCTATGGACTCTCTGGCTTGGCTCTATGGGCAATTC 180

QY 190 TCAGCGCACTCTGGCGCAGAGCTCTTCGGGTACCTGGAGCACCGGGGTGGCGG 249
    |||||
Db 181 AGCGCACACTAGTGGCAGAGCTCTTCGGTACCTGGAGCACCGAAGGGTGGCAGG 240

QY 250 GCGGCGCG-----GGGCGCGCTGGATGACGAGCCAGCCGCGCAGTGTGGCG 294
    |||||
Db 241 GCTGCGCGCGCTCTCTTGGCGAAGGGCCCTGGATGGCCACTGACGAGCGTGCA 300

QY 295 CTGACCATCTCGCTTACAGAGAGACCCCGGTACTGCGCACTGCTGCGCTCGCG 354
    |||||
Db 301 CTCACCATCTCAGCCTACCAAGAGATCCCGCTTACCTGGCGCAGTGTGACCT 360

QY 355 GCGCGCTCTGTACCGCGCGCGCTGCGCGCTCCTCATGCTGGTGGTGGAGCGCA 414
    |||||
Db 361 CCGCGCTTGTGTACCGCGCACAGAGTTACGCGCTCATGCTGGTGGAGCGCAACG 420

QY 415 GCCGAGGACCTTACATGCTGACATGTTCCGCGAGGCTTCTCGTACGAGGACCC 474
    |||||
Db 421 GCTGAGGATCTGATGCTGACATGTTCCGAGAGTCTTCCGCGATGAGGACCCG 480

QY 475 AGTACGTGTGGGACGGCACTACCAACAGCCCTGGGAACCCCGCGCGG---CG 531
    |||||
Db 481 ACTTATGTGTGGATGGCACTTACCATCAGCCCTGGGAACCAAGCGGAGGCTAC 540

QY 532 GTGGCGCGGAGCCTATCGGAGGTGGAGCGGAGGATCTCTGGGCGCTGCGAGT 591
    |||||
Db 541 GTGCGTGAAGTGCTTACCGGAGGTGGAGCGGAGGACCCCGGCGGTGGCGTGG 600

QY 592 GCGCTGTTGAGGACTCGCAGTGTGCTGCTGGCGCAGCGCTGGGCGGCAAGCGCAG 651
    |||||
Db 601 GCGCTGTTGAGAACACGCAAGTGTGCTGCTGGCTCAGCGTTGGGCGGCAAGCG 660

QY 652 GTCATGTACAGCCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGAGGTCTGT 711
    |||||
Db 661 GTCATGTACAGCTTTCAAGGCACTGGGCGCACTCCGTGGACTACGTGAGGTCT 720

QY 712 TCAGACACAGGTTGAGCCCATGCTGCTGGAGCTCTGCGGCTTCTGAGGAGGAC 771
    |||||
Db 721 TCAGACACAGAACTAGACCCCATGCACTGTCTGGAGCTTGTGCGAGTGTGTGA 780

QY 772 CCGCGGTAGGGCTGTTGGTGGGACGCTGCGGATCTTAACTCTGAGCTCTCTGG 831
    |||||
Db 781 CCGCGGTAGGGCTGTTGGAGGGGATGTGAGATCTTAACTCTTGAATCTCTGG 840

QY 832 AGCTTCTAAGCAGCTCGATACCTGCGGTAGGCTTCAATGTGGAGCGGGCTTGT 891
    |||||
```

Db 778 TCGCTGACGAGGACCCCGCCAGTACGTGT-GGACGGCACTATACCACGAGCCCTGGGAAA 720
Qy 514 CCCGGCGGGCGCGGTGGCGCCGAGCTATCGGAGGTGAGGCGGAGATCCT 573
Db 719 CCNCGNCGNCGGCGCGGTGGCGCCGAGCTATCGGAGGTGAGGCGGAGATCCT 660
Qy 574 GGGCGGCTGGCAGTGGAGCGCTGCTGAGGACTCGCAGGTGGTGGTGGCGCAGCGC 633
Db 659 GNCGGCTGGCAGTGGAGCGCTGCTGAGGACTCGCAGGTGGTGGTGGCGCAGCGC 600
Qy 634 TGGGCGGCAAGCGGAGGTCTATGACAGCTTCAAGGCGCTCGGAGATTCGGTGGAC 693
Db 599 TGGGCGGCAAGCGGAGGTCTATGACAGCTTCAAGGCGCTCGGAGATTCGGTGGAC 540
Qy 694 TACGTGCAAGTCTGACCTCGACACAAGTTGGACCCCATGGCACTCTCGAGGCTCTGT 753
Db 539 TACGTGCAAGTCTGTAAGTTCGGACACAAGTTGGACCCCATGGCACTCTCGAGGCTCTGT 480
Qy 754 CGGTGACTGGAGGAGACCCCGGTAGGGCTGTGTGGGAGCTGCGGATCCTTAAC 813
Db 479 CGGTGACTGGAGGAGACCCCGGTAGGGCTGTGTGGGAGCTGCGGATCCTTAAC 420
Qy 814 CTTCTGAGTCTCTGGGTGAGCTTCTTAAGCAGCTCGGATCTGGGTAGCCTTCAAGTGT 873
Db 419 CTTCTGAGTCTCTGGGTGAGCTTCTTAAGCAGCTCGGATCTGGGTAGCCTTCAAGTGT 360
Qy 874 GAGCGGCTTGTGAGGACTTCTCACTGTGTATCTGTCATCAGCGGTCTCTAGGCTTA 933
Db 359 GAGCGGCTTGTGAGGACTTCTCACTGTGTATCTGTCATCAGCGGTCTCTAGGCTTA 300
Qy 934 TATAGGAATAACTCTTTCAGCAGTCTTCTGAGGCTGCTGTAACAACCAAGTCTCTGGGT 993
Db 239 TATAGGAATAACTCTTTCAGCAGTCTTCTGAGGCTGCTGTAACAACCAAGTCTCTGGGT 240
Qy 994 ACCCACTGACTTTTGGGATGACCGGCACTTCAACAACCGCATGCTCAGCATGGTTAT 1053
Db 239 ACCCACTGACTTTTGGGATGACCGGCACTTCAACAACCGCATGCTCAGCATGGTTAT 180
Qy 1054 GGTACCAAGTA 1064
Db 179 GGTACCAAGTA 169

RESULT 7
AK053726
LOCUS AK053726 3542 bp mRNA linear HTC 03-APR-2004
DEFINITION Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:EI30302P16 product:hyaluronan synthase1, full insert sequence.
ACCESSION AK053726
VERSION AK053726.1 GI:26343670
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 2
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED 11076861
4
TITLE The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
AUTHORS Functional annotation of a full-length mouse cDNA collection
NATURE 409, 685-690 (2001)
5
TITLE The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
AUTHORS Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
NATURE 420, 563-573 (2002)
6 (bases 1 to 3542)
ADACHI, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.jp/
URL:http://fantom.gsc.riken.jp/.
FEATURES
source Location/Qualifiers
1. 3542
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM.DB:EI30302P16"
/db_xref="taxon:10090"
/clones="EI30302P16"
/tissue_type="eyeball"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
61. _456
/note="unnamed protein product; hyaluronan synthase1 (MGDI:106590, GB|NM_008215, evidence: BLASTN, 99%, match=762)
putative"
/codon_start=1
/protein_id="BAC35492.1"
/db_xref="GI:26343671"
/translation="MRQDMFKPSEARCCSGLARRALTIIIFALLILGLMTWVAAGVP LADRYGLLAFGLYGFALSHVAQLFAYLEHRRVAAARRSLAKGLDARTARSA LTISSAQEDPAYLRQCLTSARALLYPHTR"
ORIGIN

Query Match 45.0%; Score 479.4; DB 3; Length 3542;
Best Local Similarity 81.6%; Pred. No. 5.6e-90;
Matches 598; Conservative 0; Mismatches 116; Indels 19; Gaps 3;

QY 10 CAGGAGCGCCCAAGCCACTCTCGAGCGCCGCGCTGCTCGGCTTGGCCGCGAGGGTG 69
DB 67 CAGGACATGCCAAAGCCCTCAGAGCAGCGCGTTGCTGCTGCGCTGCGCAGCGAGCA 126
QY 70 CTGACCATCGCCTTGGCCCTGCTCATCTGGGCGCTCATGACCTGGGCGCTTACGCCCGCGG 129
DB 127 CTCAGATCATCTTGGCCCTGCTCATCTGGGCGCTCATGACCTGGGCGCTTACGCCCGAGGC 186
QY 130 GTGCGCTGGCCTCGATCGCTACCGCCTCTGGGCTTGGGCTCTACCGGCGCTTTCCTT 189
DB 187 GTTCTCTGGCTTCAGATCGCTATGCACTCTGGGCTTGGGCTCTATGGGCGATTCCTC 246
QY 190 TCAGCGCACTGTTGGCGAGAGCTCTTCGGTACCTGAGCAGCGGGGGTGGCGGG 249
DB 247 AGCGCACACCTAGTGGCACAGAGCTCTTGGCTTACCTGGAGCACCGAAGGGTGGCAGCG 306
QY 250 GCGGCGCG-----GGGCGCGCTGGATGACGACCGCGCGCGAGTGTGGCG 294
DB 307 GCTGGCGCGGCTCTTGGCGAGGGGCCCTTGGATGGGCGCACTGCGAGCGGTGGCA 366
QY 295 CTGACCATCTCGCCTACCGAGGAGCCCGGCTACCTCGCGCAGTGTGCTGGCGTCCGCC 354
DB 367 CTCACCATCTCAGCCTACCAAGAGGATCCCGCTTACCTGGCGCAGTGTGACCTCCGG 426
QY 355 CGCGCCCTGCTGTACCGCGCGCGCTGGCGCTCTCATGTTGGTGGATGGCAACCGC 414
DB 427 CGCGCCTTGTGTACCGCGCACAGAGGTAACGCGGTCTCATGTTGGTGGACGCGCAACCGC 486
QY 415 GCGGAGGACCTCTACATGCTGCATGTTTCGCGAGGCTCTTCGCTGACGAGGACCGCGCC 474
DB 487 GCTGAGGATCTGATGTTGGTGCATGTTCCGAGAAGTCTTCGCGCATGAGGACCGCGCC 546
QY 475 ACCTATGTTGGATGGCACTTACCATCAGCCCTGGGAACCGAGCGGTACGCGCGCT 606
QY 532 GTGGCGCGCGAGCTATCGGAGGTGAGGGGAGGATCTTCGGCGCGCTGCGAGTGGAG 591
DB 607 GTCGTTGAAGGTGCTTACCGGAGGTGGAGCGGAGGACCGCGCGCGTGGCGGTGAG 666
QY 592 CGCTGTGTAGAGAC--TCGACGTGCTGTGGTGGCGCAGCGCTGGGCGCGCAACGCGGA 650
DB 667 GCGTGTGTAGAACGACGCGAGTGGTGTGGTGGCTCAGCGTTGGGCGCGCAACGTTGA 726
QY 651 GGTCACTGTACAGCCTTCAAGCGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGTGA 710
DB 727 GGTCACTGTACAGCCTTTCAGGCACTGGGCGACTCCGCTGCACTACGTGCAAGTGTGA 786
QY 711 CTCGACACAAAGG 723
DB 787 TGGGTTTCCAGG 799

RESULT 8
BM544718
LOCUS
DEFINITION AGENCOURT 6494603 NIH_MGC_124 Homo sapiens cdna clone IMAGE:5727416
5', mRNA sequence.
ACCESSION BM544718
VERSION BM544718.1 GI:18776197
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1084)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM12721 row: i column: 15
High quality sequence start: 5
High quality sequence stop: 689.
Location/Qualifiers
1. 1084
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5727416"
/issue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_124"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: EcoRV
(destroyed); Site_2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."

FEATURES

source

ORIGIN

Query Match 41.5%; Score 442.4; DB 4; Length 1084;
Best Local Similarity 99.8%; Pred. No. 3e-82; Mismatches 0; Indels 0; Gaps 0;
Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 621 CGTGGCGCAGCGCTGGGCGCGCAAGCGCGAGTCTATGTACACAGCTTCAAGCGCTCGG 680
DB 20 CGTGGCGCAGCGCTGGGCGCGTAAGCGCGAGTCTATGTACACAGCTTCAAGCGCTCGG 79
QY 681 AGATTCCGTGACATACGTGCAGGTCTGTGACTCGGACACAAGTTGGACCCCATGGCACT 740
DB 80 AGATTCCGTGACATACGTGCAGGTCTGTGACTCGGACACAAGTTGGACCCCATGGCACT 139
QY 741 GCTGAGCTCTGCGGGTACTGGACGAGGACCCCGGGTAGGGGCTGTGTGGGGACGT 199
DB 140 GCTGAGCTCTGCGGGTACTGGACGAGGACCCCGGGTAGGGGCTGTGTGGGGACGT 199
QY 801 GCGGATCCTTAACCCCTCTGGACTCTGGGTCAGCTTCTTAAGCAGCTTCCGATACTGGGT 860
DB 200 GCGGATCCTTAACCCCTCTGGACTCTGGGTCAGCTTCTTAAGCAGCTTCCGATACTGGGT 259
QY 861 AGCCTTCAATGTGAGCGGGCTTGTGAGACTATTTCACATGTGTATCTGCATCAGCGG 920
DB 260 AGCCTTCAATGTGAGCGGGCTTGTGAGACTATTTCACATGTGTATCTGCATCAGCGG 319
QY 921 TCCTCTAGGCTATATAGGATACCTCTTGCACGAGTCTTTCAGCGCTTGGTCAACCA 980
DB 320 TCCTCTAGGCTATATAGGATACCTCTTGCACGAGTCTTTCAGGCTTGGTCAACCA 379
QY 981 GAAGTCTCTGGTACCCACCTGTACTTTTGGGGATGACCGGACCTTCAACCAACCGCATGCT 1040
DB 380 GAAGTCTCTGGTACCCACCTGTACTTTTGGGGATGACCGGACCTTCAACCAACCGCATGCT 439
QY 1041 CAGCATGGGTATGCTTACCAAGTA 1064
DB 440 CAGCATGGGTATGCTTACCAAGTA 463

RESULT 9

BG707429

LOCUS

DEFINITION

BG707429 551 bp mRNA linear EST 07-MAY-2001
602672848F1 NIH_MGC_96 Homo sapiens cdna clone IMAGE:4795643 5',


```

mRNA sequence.
BG707429      1  GI:13983769
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 551)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10678 row: p column: 12
High quality sequence stop: 551.
FEATURES             source
    Location/Qualifiers
        1..551
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4795643"
        /tissue_type="hypothalamus"
        /lab_host="DH10B"
        /clone_lib="NIH_MGC_96"
        /note="Organ: brain; Vector: pBluescriptR (modified
        pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
        (gtcgag); Oligo-dT primed using primer
        5'-TTTTTTTTTTTTTTVN-3', size-selected for average
        insert size 2.3 kb and normalized to ROT 5. This is a
        primary library enriched for full-length clones and
        constructed using the Cap-trapper method (Carninci, in
        preparation). Library constructed by M. Brownstein
        (NIMH/NHGRI, National Institutes of Health). Note: this is
        a NIH_MGC Library."
ORIGIN
Query Match      40.2%; Score 428.4; DB 4; Length 551;
Best Local Similarity 99.8%; Pred. No. 2.5e-79;
Matches 429; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

635  GGGGGCGCAGCGGAGGTGATGTACACAGCCTTCAAGCGCTCGGAGATTGGTGACT 694
      |||||||
Db    6  GGGGGCGCAGCGGAGGTGATGTACACAGCCTTCAAGCGCTCGGAGATTGGTGACT 65

695  ACGTGCAGGTCGTGACTCGGACACAAGGTTGGACCCCATGGCACTGCTGGAGCTCGTGC 754
      |||||||
Db    66  ACGTGCAGGTCGTGACTCGGACACAAGGTTGGACCCCATGGCACTGCTGGAGCTCGTGC 125

755  GGGTACTGGACGAGACCCCGGTAGGGCTGTTGGTGGGACGTGGGATCCTTAACC 814
      |||||||
Db    126  GGGTACTGGACGAGACCCCGGTAGGGCTGTTGGTGGGACGTGGGATCCTTAACC 185

815  CTCTGGACTCTGGGTGAGCTTCTTAAGACGCTTGCATGCTGAGTGGTGGTCAATCTGG 874
      |||||||
Db    186  CTCTGGACTCTGGGTGAGCTTCTTAAGACGCTTGCATGCTGAGTGGTGGTCAATCTGG 245

875  AGCGGGCTTGTTCAGAGTACTTCCACTGTGTATCCTGCATCAGCGGTCTCTAGGCTAT 934
      |||||||
Db    246  AGCGGGCTTGTTCAGAGTACTTCCACTGTGTATCCTGCATCAGCGGTCTCTAGGCTAT 305

935  ATAGGAATAACCTTTTGACGAGCTTTCTTAGGCTCTGGTACAACAGAGTTCCTGGGTA 994
      |||||||
Db    306  ATAGGAATAACCTTTTGACGAGCTTTCTTAGGCTCTGGTACAACAGAGTTCCTGGGTA 365

995  CCCACTGTACTTTTGGGATGACCGGCACCTCACCACCGCATGCTCAGCATGGTTATG 1054
      |||||||
Db    366  CCCACTGTACTTTTGGGATGACCGGCACCTCACCACCGCATGCTCAGCATGGTTATG 425

1055  CTACCAAGTA 1064
      |||||||
Db    426  CTACCAAGTA 435

RESULT 10
BG707608      1  GI:13984125
LOCUS
DEFINITION
mRNA sequence.
BG707608      1  GI:13984125
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 551)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: csapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10673 row: o column: 17
High quality sequence stop: 551.
FEATURES             source
    Location/Qualifiers
        1..551
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:4793704"
        /tissue_type="hypothalamus"
        /lab_host="DH10B"
        /clone_lib="NIH_MGC_96"
        /note="Organ: brain; Vector: pBluescriptR (modified
        pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
        (gtcgag); Oligo-dT primed using primer
        5'-TTTTTTTTTTTTTTVN-3', size-selected for average
        insert size 2.3 kb and normalized to ROT 5. This is a
        primary library enriched for full-length clones and
        constructed using the Cap-trapper method (Carninci, in
        preparation). Library constructed by M. Brownstein
        (NIMH/NHGRI, National Institutes of Health). Note: this is
        a NIH_MGC Library."
ORIGIN
Query Match      40.2%; Score 428.4; DB 4; Length 551;
Best Local Similarity 99.8%; Pred. No. 2.5e-79;
Matches 429; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

635  GGGGGCGCAGCGGAGGTGATGTACACAGCCTTCAAGCGCTCGGAGATTGGTGACT 694
      |||||||
Db    6  GGGGGCGCAGCGGAGGTGATGTACACAGCCTTCAAGCGCTCGGAGATTGGTGACT 65

695  ACGTGCAGGTCGTGACTCGGACACAAGGTTGGACCCCATGGCACTGCTGGAGCTCGTGC 754
      |||||||
Db    66  ACGTGCAGGTCGTGACTCGGACACAAGGTTGGACCCCATGGCACTGCTGGAGCTCGTGC 125

755  GGGTACTGGACGAGACCCCGGTAGGGCTGTTGGTGGGACGTGGGATCCTTAACC 814
      |||||||
Db    126  GGGTACTGGACGAGACCCCGGTAGGGCTGTTGGTGGGACGTGGGATCCTTAACC 185

```

Qy	815	CTCTGGACTCTCTGGGTGAGCTTCTTAAGCAGCCTGCAGTACTTGGGTAGCCTTCAATGTGG	874
Db	186	CTCTGGACTCTCTGGGTGAGCTTCTTAAGCAGCCTGCAGTACTTGGGTAGCCTTCAATGTGG	245
Qy	875	AGCGGGCTTGTGCAGAGTACTTCCACATGTGTATCCTGCATCAGCGGTCTCTTAGGCCCTAT	934
Db	246	AGCGGGCTTGTGCAGAGTACTTCCACATGTGTATCCTGCATCAGCGGTCTCTTAGGCCCTAT	305
Qy	935	ATAGGAATAACCTCTTCGAGCAGTTTCTTGAGGCTGTGTACAAACGAAAGTTCTCGGGTA	994
Db	306	ATAGGAATAACCTCTTCGAGCAGTTTCTTGAGGCTGTGTACAAACGAAAGTTCTCGGGTA	365
Qy	995	CCCACTGTACTTTTGGGGATGACCGGCACCTCACCAACCGCATCTCAGCATGGGTTATG	1054
Db	366	CCCACTGTACTTTTGGGGATGACCGGCACCTCACCAACCGCATCTCAGCATGGGTTATG	425
Qy	1055	CTACCAAGTA	1064
Db	426	CTACCAAGTA	435

RESULT 11	
BX327795	
LOCUS	610 bp mRNA linear EST 08-APR-2004
DEFINITION	BX327795 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CSODI027YJ02 5-PRIME, mRNA sequence.
ACCESSION	BX327795
VERSION	BX327795.2 GI:462777978
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 610) Li,W.B., Gruber,C., Jesses,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished (2001) On May 2, 2003 this sequence version replaced gi:30340686.
AUTHORS	
TITLE	
JOURNAL	
COMMENT	

Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Crémieux, Cp 5706 - 91057 EVRY cedex - FRANCE
E-mail: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
The NotI strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9502.f

For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?s=CS0BAG028ZH05> CS02640 1&c=9502.1

```

FEATURES
source
Location/Qualifiers
1. 610
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1027VJ02"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dt)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

```

Query Match	37.1%	Score 395.4;	DB 5;	Length 610;
Best Local Similarity	98.3%	Pred. No. 2.1e-72;		
Matches 399;	Conservative	0;	Mismatches 7;	Indels 0;
				Gaps 0;

Qy 659 ACACAGCCTTCAAGGCGCTCGGAGATTTCGGTGGACTACGTGCAAGTCTGTGATCGGACA 718
|||
Db 1 ACATGCTCTTAAGGNCGCTCGGAGATTTCGGTGGACTACGTGCAAGTCTGTGATCGGACA 60
|||

Qy	719	CAAGGTTGGACCCCATGGCACTGCTGGAGCTCGTGGGGTACTGGACGAGGACCCCGGG	778
Db	61	CAAGGTTGGACCCCATGGCACTGCTGGAGCTCGTGGGGTACTGGACGAGGACCCCGGG	120
Qy	779	TAGGGGCTGTTGGTGGGGACGTGGGATCCTTAACCTCTGGACTCCTGGGTGAGCTTCC	838
Db	121	TAGGGGCTGTTGGTGGGGACGTGGGATCCTTAACCTCTGGACTCCTGGGTGAGCTTCC	180
Qy	839	TAAGCAGCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCC	898
Db	181	TAAGCAGCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGCTACTTCC	240
Qy	899	ACTGTGTATCCTGCATCAGCGGTCTCTAGGCCCTATATAGGAATAA	958
Db	241	ACTGTGTATCCTGCATCAGCGGTCTCTAGGCCCTATATAGGAATAA	300
Qy	959	TTCTTGAGGCTGGTACACCAAGGTTCTGGGTACCACTGACTCTTTTGGGGATGACC	1018
Db	301	TTCTTGAGGCTGGTACACCAAGTTCCTGGGTACCACTGACTCTTTTGGGGATGACC	360
Qy	1019	GGCACCTCACCAACCGCATGCTCAGCATGGGTTATGCTACCAAGTA	1064
Db	361	GGCACCTCACCAACCGCATGCTCAGCATGGGTTATGCTACCAAGTA	406

RESULT 12					
BI753116					
LOCUS	BI753116	782 bp	mRNA	linear	EST 25-SEP-2001
DEFINITION	603025961F1 NIH_MGC_114 Homo sapiens CDNA clone IMAGE:5196399 5', mRNA sequence.				

ACCESSION	BI753116
VERSION	BI753116.1
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1.	(Pages 1 to 782)	NIH-MGC http://mgc.nci.nih.gov/ . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)		mammalian, eutherian, primates, cetartiodactyla, hominidae, homo.
	Contact: Robert Strausberg, Ph.D. Email: cgapsb@mail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov			

```

FEATURES
  source      1. .782
              Location/Qualifiers
High quality sequence stop: 780.
Plate: LLAM11491 row: j column: 16
ncpp://image.llnl.gov

```

```

I. .782
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5196399"
/lab_host="DH10B"
/clone_lib="NIH_MGC_114"
/notes="EcorV; Brain; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: EcorV (destroyed); RNA source anonymous pool of 6
male brains, age range 23-27 yo. Library is oligo-dT
primed and directionally cloned (EcorV site is destroyed
upon cloning). Average insert size 1.5 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(invitrogen). Research Genetics tracking code 019. Note:
this is a NIH MGC Library."

```

Query Match 35.7%; Score 380.4; DB 4; Length 782;

```
Best Local Similarity 99.5%; Pred. No. 2.8e-69;
Matches 392; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 671 AGGCGCTCGGAGATTTCGGTGGACTACGTGAGGTTCTGTGACTGCGACACAAGTTGGACC 730
Db 6 AAGCGCTCGGAGATTTCGGTGGACTACGTGAGGTTCTGTGACTGCGACACAAGTTGGACC 65
QY 731 CCATGGCACTGCTGGAGCTCGTGGCGGTACTGAGCAGGAGACCCCGGTAGGGGCTGTTG 790
Db 66 CCATGGCACTGCTGGAGCTCGTGGCGGTACTGAGCAGGAGACCCCGGTAGGGGCTGTTG 125
QY 791 GTGGGGAGCTGCGGACTCTTAACCTCTTGGACTCTCGGCTGAGCTTCTTAAGCAGCCTGCG 850
Db 126 GTGGGGAGCTGCGGACTCTTAACCTCTTGGACTCTCGGCTGAGCTTCTTAAGCAGCCTGCG 185
QY 851 GATACCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTTCACCTGTGTATCCT 910
Db 186 GATACCTGGGTAGCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTTCACCTGTGTATCCT 245
QY 911 GCATCAGCGGTCTTAGGCTTATAGGATTAACCTCTTGCAGCAGTCTTTCAGGCGCT 970
Db 246 GCATCAGCGGTCTTAGGCTTATAGGATTAACCTCTTGCAGCAGTCTTTCAGGCGCT 305
QY 971 GGTACCAACGAGAGTTCTTGGGTACCCACTGTACTTTTGGGATGACCGGACCTCACCA 1030
Db 306 GGTACCAACGAGAGTTCTTGGGTACCCACTGTACTTTTGGGATGACCGGACCTCACCA 364
QY 1031 ACCGATGCTCAGCATGGTTATGTTACCAAGTA 1064
Db 365 ACCGATGCTCAGCATGGTTATGTTACCAAGTA 398

RESULT 13
BB665418
LOCUS
DEFINITION
BB665418 RIKEN full-length enriched, 0 day neonate eyeball Mus
musculus cDNA clone E130302P16 5', mRNA sequence.
ACCESSION
BB665418
VERSION
BB665418.1 GI:16499056
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 650)
AUTHORS
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T.,
Matsunura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.
and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
```

```
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers
1. 650
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="E130302P16"
/tissue_type="eyeball"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 0 day neonate
eyeball"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGCGCGCGCACTCGAGTTTCTTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGAGATTCTCGAGTTAATTAATTAATCCCGCCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from Lambda FLC I."
ORIGIN
Query Match 33.6%; Score 357.4; DB 2; Length 650;
Best Local Similarity 77.4%; Pred. No. 1.8e-64;
Matches 452; Conservative 0; Mismatches 117; Indels 15; Gaps 1;
QY 10 CAGGACGCGCCCAAGCCACTCTCTGACCGCCGCGCTGCTCCGCGCTGGCCCGGAGGTG 69
Db 67 CAGGACATGCCAAAGCCCTCAGAGGCGAGCGGTTGCTGCTCTGGCCTGGCCGAGCGAGCA 126
QY 70 CTGACCATCGCTTCGCGCTGCTCATCTCTGGGCTCATGACCTGGGCTAGCGCCGCGGG 129
Db 127 CTCAGCATCATCTTTGCGCTGCTCATCTGCGGCTCATGACCTGGGCTAGCGCCGAGGC 186
QY 130 GTGCGCTGCGCTCCGATCGCTAGCGCTCTCTGCGCTTCTGCGCTCTAGCGGCGCTTCTT 189
Db 187 GTTCTCTGCGCTTCAGATCGCTATGACTCTCTGCGCTTCTGCGCTCTATGCGGCAATTC 246
QY 190 TCAGCGCACCTGTGGCGCAGAGCCTCTTCGCGTACTCTGGAGCACCGCGCGGTGGCGCG 249
Db 247 AGCGCACACCTAGTGGCACAGAGCCTCTTCGCTTACCTGAGGACACCGAAGGTTGGCAGCG 306
QY 250 GCGGCGCG-----GGGCGCGCTGATGACGCCACCGCGCGGCTGCGGCTGCGCG 294
Db 307 GCTGCGCGCGCTCTTCTTGGCGAAGGGGCGCCCTGATGCGGCCACTGACGCGAGCGTGCA 366
QY 295 CTGACCATCTCCGCTTACCAAGGAGGACCGCGCTGCTGCGCGCAGTGCCTGGCGCTCGGCC 354
Db 367 CTACCATCTCAGCCTACCAAGAGGATCCGCTTACCTGCGCCAGCTGCTTGACCTCGCG 426
QY 355 CGCGCGCTCTGTATCCGCGCGCGCGCTCGCGCTCTCATGTTGTTGATGCGCAACCGC 414
Db 427 CGCGCGCTTGTGTATCCGCGCACAGAGGTAACGGGTGCTCATGTGTGTGGAGCGCAACCGC 486
```

```
QY 415 GCGGAGGACCTTACATGTCGACATGTTCCGCGAGGTCTTCCTGACGAGGACCCGCC 474
|||
Db 487 GCTGAGGATCTGTACATGTCGACATGTTCCGAGAAGTCTTCGCCGATGAGGACCCGCC 546
|||
QY 475 AGCTAGTGTGGGAGGCGCACTACACACCGCTGGGAACCCGCCGCGCGCGCGGTG 534
|||
Db 547 ACTTATGTGTGGATGGCACTACCATGACCCCTGGGACCGGGAAGCTACAGGCGCTG 606
|||
QY 535 GCGCGCGGAGCCTATCGGAGGTGGAGCGGAGGATCCTGGGCG 578
|||
Db 607 TCGGTGAAGGCGCTACCGGAGGTGAGCGAGAGACCCCGGCG 650
|||

RESULT 14
LOCUS AY421570
DEFINITION Pan troglodytes HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY421570
VERSION AY421570.1 GI:39748429
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1 (bases 1 to 739)
AUTHORS Clark,A.G., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 739)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..739
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>739
/gene="HAS1"
/locus_tag="HCM7601"
ORIGIN
Query Match 29.7%; Score 315.8; DB 9; Length 739;
Best Local Similarity 99.4%; Pred. No. 9.4e-56;
Matches 317; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 746 AGCTGTGGGGTACTGACGAGGACCCCGGGTAGGGCTTTGGTGGGACGTGCGGA 805
|||
Db 1 AGCTGTGGGGTACTGACGAGGACCCCGGGTAGGGCTTTGGTGGGACGTGCGGA 60
|||
QY 806 TCCTTAACCTCTGACCTCTGGTCAGCTTCTTAAGCAGCCTCGATCTGGGTAGCCT 865
|||
Db 61 TCCTTAACCTCTGACCTCTGGTCAGCTTCTTAAGCAGCCTCGATCTGGGTAGCCT 120
|||
QY 866 TCAATGTGGAGGGGCTTGTTCAGAGCTACTTCCACTGTGTATCTTCGCATCAGCGTCTTC 925
|||
Db 121 TCAATGTGGAGGGGCTTGTTCAGAGCTACTTCCACTGTGTATCTTCGCATCAGCGTCTTC 180
|||
QY 926 TAGGCCTATATAGGAATACCTCTTGGCAGCAGTTCTTGGAGCGCTGGTACACCAAGAGT 985
|||
```

```
Db 181 TAGGCCTCTATAGGAATACCTCTTGGCAGCAGTTCTTGGAGCGCTGGTACACCAAGAGT 240
|||
QY 986 TCCTGGGTACCCACCTGTACTTTTGGGGATGACCGCACCTACCAACCGCATGCTCAGCA 1045
|||
Db 241 TCCTGGGACCCACCTGTACTTTTGGGGATGACCGCACCTACCAACCGCATGCTCAGCA 300
|||
QY 1046 TGGGTTATGCTACCAAGTA 1064
|||
Db 301 TGGGTTATGCTACCAAGTA 319
|||

RESULT 15
LOCUS AK028582
DEFINITION Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched
library, clone:4732404L04 product:similar to DG42III [Brachydanio
rerio], full insert sequence.
ACCESSION AK028582
VERSION AK028582.1 GI:26324525
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
PUBMED 20499374
MEDLINE 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yanamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
PUBMED 20530913
MEDLINE 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
REFERENCE 6
AUTHORS Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hasegaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
```

Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in RIKEN.
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

FEATURES

Location/Qualifiers

1..4151
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CS7BL/6J"
/db_xref="PANTOM_DB:4732404L04"
/db_xref="taxon:10090"
/clone="4732404L04"
/tissue_type="skin"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
118..1782

CDS

/note="unnamed protein product; putative
similar to DG42I1 (Brachydanio rerio) (SPTR|Q9DG40,
evidence: FASTA, 75.1%ID, 99.8%length, match=1659)"
/codon_start=1
/protein_id="BAC26017.1"
/db_xref="GI:26324526"
/translation="MPQLTTLARVGTSLPALVGLLAAYVGYQFIHTEKHVLS
FGLGAILGLHLLQSLFAPLEHRRMRAPGLKHCQRPRFVALCIAAYQEDPEYL
RKLRSAQRIAPFLNKKVNVVDGQEDTYMLDIPHEVLGGTEAGFGFVWRNSPEAG
EGETEASIQEMERVRVAVMASTFCINQMRGRVEMYTAFKALGNSVDYIQVCDSD
TVLPDPACTIEMLRVLEEDPQVGGVDQILNKYDSWISFLSSVRYMAFNVERACOS
YFGVCQISGPLGMYRNSLQOFLQEDWYHOKFLGSKCSFGDDRLHTRNLSLIGVRYKY
TARSKCTETPTVRLWLNQOTRWKSVFRELWNSLWFKHKLHWMTVESVYGFPPF
FLIATVQLFRGIRWILLFLLVQLVGLIKATYACPLRGNAXEIMFSLVSLYMS
LLPAKIFAIATIRKSGWTSGRKTIIVNFIPLVSIWVAVLLGLLAYTAYCQDLFSE
TELAPLVSGALYGCYWVALLMLYLAIIRRCGKPKBQSYSLAFV"
4131..4136
/note="putative"
4151
/note="putative"

polyA_signal

polyA_site

ORIGIN

Query Match 29.3%; Score 311.6; DB 3; Length 4151;
Best Local Similarity 60.3%; Pred. No. 7.4e-55;
Matches 609; Conservative 0; Mismatches 359; Indels 42; Gaps 4;
Qy 61 CGAGGGGTGACCATCGCCTTGGCTTCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
Db 145 CGTGTGGTGGGACCAAGTCGTCTTTGGCTTGGTAGTGTGGGAGGCATCTCTGGGCTTAT 204
Qy 121 GCCCGGGGTGGCGCTCGATCGCTACGGCTCTCGGCTTGGCTTGGCTTCTACGGG 180
Db 205 GTGACAGGCTACCAAGTTTATCCACAGAAAGCACTACCTGTCTTTGGCTTCTACGGT 264
Qy 181 GCCTTCTCTTTCAGGCGACCTGCTGGTGGCGAGAGCCTCTTTCGGGTACCTGGAGCACCGGCGG 240
Db 265 GCCATCTGGGTCTACATCTGCTCATCCAGAGCCTGTTTGGCTTCTCTGGAGCAACGTGA 324
Qy 241 GTGGCGGCGGCGGCGCG-----GGGGCGGCTGTGATGAGCCACCGCGCGCAGTGTGGCG 294
Db 325 ATGGCGAGGGCGGCGGCGCGCTCAAGCTGCACTGCTCCAGAGCGCGCGCTTCAGTGGCA 384
Qy 295 CTGACCATCTCGGCTACCAAGGAGGAGCCCGCGTACCTCGGCCAGTGCTGGCGTCCGCC 354

Db 385 CTCTGCATTGCTCTACCAAGAGGACCCCGAATACTGCGCAAGTGCTTCGCTCAGCT 444
Qy 355 CGCGCCCTGCTGTATACCGCGCGCGGCTGCGCTCTCATGTGGTGGATGGCAACCGC 414
Db 445 CAGCGCATGGCTCTTCCAAAC-----CTCAAGGTGGTCTATGGTAGTGGATGGCAATCGC 498
Qy 415 GCCGAGGACCTCTATAGTGTGCAGATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCC 474
Db 499 CAGGAAGATACCTATGTTGGACATCTTCATGAGGTGCTGGGTGGCCTGAGCAAGCT 558
Qy 475 ACGTACGTGTGGGACGCAACTACACACAGCCCTGGCAACCCGCGCGCGCGCGCGGTG 534
Db 559 GGCTCTTGTGTGGCGTAGCAATTCCA-----TGAGCGGGGTGAAGGA 603
Qy 535 GGCGCCGAGCCTATACGGGAGGTGGAGCGGAGAGATCCTGGGCGGCTGGCAGTGGAGGCG 594
Db 604 GAGACAGAGGCGCAGCCTGCAGGAAGGCATGGAGCGT-----GTGCGAGCT 648
Qy 595 CTGCTGAGGACTCGCAGGTGGGTGTGCGTGGCGCAGCGCTGGGGGGCGGCGGAGGTC 654
Db 649 GTGGTGTGGGCGCAGCCTTCTCATGTCATCATGTCAGAAAGTGGGGGGCGAAGCGTGAGGTC 708
Qy 655 ATGTACACAGCCTTCAAGGCGCTCGGAGATTTCGGTGGACTACGTGCAGGTCTGTGACTCG 714
Db 709 ATGTACACTGCTTCAAGGCCCTTGGCAACTCAGTGGACTACATCCAGGTGTGTGACTCT 768
Qy 715 GACAAAGTGTGGACCCCATGGCACTGCTGGAGCTCTGCGGGTACTGGAGAGGAGACCC 774
Db 769 GACACTGTCTGGACCCAGCCTGCACCATTTGAGATGCTTCGAGTCTTGGAAAGATCCC 828
Qy 775 CGGCTAGGGCTGTGGTGGGAGCTCGGATCCTTAACCTCTGGACTCTGGGTCAGC 834
Db 829 CAAGTAGGAGGTGTGGAGGAGATGTCCAAATCCTCAAGTATGATTATGATCTTC 888
Qy 835 TTCTTAAGCAGCCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGCTTGTGACAGGTAC 894
Db 889 TTCTTGAGCAGTGTGAGGTACTGGATGCTTCAACGTGGAGCGGCTGCGCACTCTTAC 948
Qy 895 TTCCACTGTGTATCCTGCATCAGCGGTCTCTAGGCTTATAGGAATAACCTCTTGCAG 954
Db 949 TTTGGCTGTGTGCAATGTATTAGTGGGCTTTGGGCATGTACCGCAACAGCCTCTCTTCA 1008
Qy 955 CAGTTTCTTGGGCTGCTACACAGAGTTCCTGGGTACCCACTGTACTTTTGGGGAT 1014
Db 1009 CAGTTCTTGGAGGATTTGTACCATCAGAAATTCCTAGGAGCAAGTGACGCTTTGGGGAT 1068
Qy 1015 GACCGCACCTTCACCAACCGCATGCTCAGCATGGGTTTATGCTACCAAGTA 1064
Db 1069 GATCGGCACCTTACCAACCGAGTCTCTGAGTCTTGGCTACCGGACTAAGTA 1118

Search completed: March 13, 2005, 11:55:00
Job time : 4393.21 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:01:16 ; Search time 209.778 Seconds
(without alignments)
8307.054 Million cell updates/sec

Title: US-10-672-399-7
Perfect score: 1065
Sequence: 1 atgagacagcaggacgcgc.....tgggtatgtaccacagtaa 1065

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1064	99.9	2116	3	US-08-635-552A-1 Sequence 1, Appli
2	1060.8	99.6	2087	4	US-09-949-016-5659 Sequence 5659, Ap
3	1051.6	98.7	2117	3	US-09-155-768-1 Sequence 1, Appli
4	697.2	65.5	14862	4	US-09-949-016-17401 Sequence 17401, A
5	396.4	37.2	601	4	US-09-949-016-201775 Sequence 201775,
6	353	33.1	1662	4	US-09-949-016-3728 Sequence 3728, Ap
7	311.6	28.3	1665	4	US-08-812-008-31 Sequence 31, Appli
8	246.8	23.2	2947	4	US-08-675-499A-1 Sequence 1, Appli
9	246.8	23.2	2947	4	US-08-812-008-1 Sequence 1, Appli
10	230.4	21.6	601	4	US-09-949-016-201780 Sequence 201780,
11	221	20.8	3003	4	US-09-949-016-737 Sequence 737, App
12	167.6	15.7	9871	4	US-09-949-016-15470 Sequence 15470, A
13	137.4	12.9	601	4	US-09-949-016-132199 Sequence 132199,
14	127.8	12.0	32176	4	US-09-949-016-12479 Sequence 12479, A
15	95.6	9.0	235	4	US-08-675-499A-25 Sequence 25, Appli
16	95.6	9.0	235	4	US-08-812-008-25 Sequence 25, Appli
17	92.4	8.7	235	4	US-08-675-499A-26 Sequence 26, Appli
18	92.4	8.7	235	4	US-08-812-008-26 Sequence 26, Appli
19	88.2	8.3	235	4	US-08-675-499A-23 Sequence 23, Appli
20	88.2	8.3	235	4	US-08-812-008-23 Sequence 23, Appli
21	82.8	7.8	6645	4	US-09-902-540-8819 Sequence 8819, Ap
22	82.8	7.8	10210	4	US-09-902-540-938 Sequence 938, App
23	69.2	6.5	1320	4	US-09-724-797-85 Sequence 85, Appli
24	66.6	6.3	3957	4	US-10-237-551-193 Sequence 193, App
25	66.6	6.3	154746	4	US-09-827-688-8 Sequence 8, Appli
26	66.6	6.3	154746	4	US-09-827-688-8 Sequence 8, Appli
27	66.2	6.2	1614	4	US-09-616-289-45 Sequence 45, Appli

c	28	66.2	6.2	12425	4	US-09-616-289-50	Sequence 50, Appli
c	29	64.8	6.1	536165	4	US-09-214-808-1	Sequence 1, Appli
c	30	62	5.8	1292	3	US-08-483-533-37	Sequence 37, Appli
	31	62	5.8	1292	3	US-09-283-471A-37	Sequence 37, Appli
	32	60.6	5.7	4257	2	US-08-690-473-1	Sequence 1, Appli
	33	60.6	5.7	4257	3	US-09-259-821A-1	Sequence 1, Appli
	34	60.6	5.7	4257	3	US-08-843-659-1	Sequence 1, Appli
	35	60.6	5.7	4257	4	US-09-825-288A-1	Sequence 1, Appli
c	36	60.2	5.7	12001	1	US-08-458-568A-11	Sequence 11, Appli
	37	60.2	5.7	2900	1	US-08-034-650-9	Sequence 9, Appli
	38	60.2	5.7	2900	1	US-08-449-015-9	Sequence 9, Appli
	39	59.6	5.6	1647	4	US-09-902-540-4211	Sequence 4211, Ap
c	40	59.6	5.6	21330	4	US-09-902-540-1209	Sequence 1209, Ap
c	41	59.4	5.6	1926	3	US-09-249-585A-4	Sequence 4, Appli
c	42	59.4	5.6	1931	2	US-09-130-114-2	Sequence 2, Appli
	43	59.4	5.6	4411529	3	US-09-103-840A-1	Sequence 1, Appli
	44	59.2	5.6	1051	2	US-08-865-273-1	Sequence 1, Appli
	45	59.2	5.6	1051	3	US-09-385-174-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-635-552A-1
; Sequence 1, Application US/08635552A
; Patent No. 6423514
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; TITLE OF INVENTION: Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millicia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36..1769
US-08-635-552A-1

Query Match 99.9%; Score 1064; DB 3; Length 2116;
Best Local Similarity 100.0%; Pred. No. 1.1e-203;
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGAGACAGCAGGACGCCCAAGCCCACTCCTGACGCCCGCTGCTCGGCTGGCC 60
|||||


```
Db 36 ATGAGACAGCAGGACGGCCCAAGCCCACTCTCTGACGCGCGCGCTGCTCGGCGCTGGCC 95
Qy 61 CGGAGGGTGTGACCATGCGCTTGGCCCTGCTCATCTCTGGGCTCATGACCTGGGCGCTAC 120
Db 96 CGGAGGGTGTGACCATGCGCTTGGCCCTGCTCATCTCTGGGCTCATGACCTGGGCGCTAC 155
Qy 121 GCGCGCGGGTGGCGCTGCGCTCCGATGCTAGCGGCTCTGGGCTCTGGGCTCTACGGG 180
Db 156 GCGCGCGGGTGGCGCTGCGCTCCGATGCTAGCGGCTCTGGGCTCTGGGCTCTACGGG 215
Qy 181 GCGCTCTCTTTTACGCGACCTGGTGGCGAGCGCTCTTGGCGTACCTGGAGCACCGGCGG 240
Db 216 GCGCTCTCTTTTACGCGACCTGGTGGCGAGCGCTCTTGGCGTACCTGGAGCACCGGCGG 275
Qy 241 GTGGCGGCGGCGCGCGGGGCGGCTGATGAGCGCACCGCGCGCAGTGTGGCGCTGAGC 300
Db 276 GTGGCGGCGGCGCGCGGGGCGGCTGATGAGCGCACCGCGCGCAGTGTGGCGCTGAGC 335
Qy 301 ATCTCCGCTTACAGGAGACCCCGGTACCTGCGCCAGTGCCTGGCGTCCGCGCGCGCC 360
Db 336 ATCTCCGCTTACAGGAGACCCCGGTACCTGCGCCAGTGCCTGGCGTCCGCGCGCGCC 395
Qy 361 CTGCTGTATACCGCGCGCGGCTGGCGCTCTCATGCTGGTGGATGGCAACCGCGCGAG 420
Db 396 CTGCTGTATACCGCGCGCGGCTGGCGCTCTCATGCTGGTGGATGGCAACCGCGCGAG 455
Qy 421 GACCTCTACATGCTGCAATGCTTGGCGTCTTGGCGTACCTGGCGTCTGGCGCTCTACGG 480
Db 456 GACCTCTACATGCTGCAATGCTTGGCGTCTTGGCGTACCTGGCGTCTGGCGCTCTACGG 495
Qy 481 GTGGGACGGCAACTACCAAGCGCTTGGGAAACCGCGCGCGGGCGCGGCTGGGCGCC 540
Db 516 GTGGGACGGCAACTACCAAGCGCTTGGGAAACCGCGCGCGGGCGCGGCTGGGCGCC 575
Qy 541 GGAGCGCTATCGGAGGTGGAGGCGGAGGATCCTTGGGCGGCTGGCGAGTGGAGCGCTGGTG 600
Db 576 GGAGCGCTATCGGAGGTGGAGGCGGAGGATCCTTGGGCGGCTGGCGAGTGGAGCGCTGGTG 635
Qy 601 AGGACTCGAGGTGGTGTGGTGGCGAGCGCTGGGCGGGGCGGAGCGGAGGTCTATGTAC 660
Db 636 AGGACTCGAGGTGGTGTGGTGGCGAGCGCTGGGCGGGGCGGAGCGGAGGTCTATGTAC 695
Qy 661 ACAGCGCTTCAAGCGCTCGGAGATTCGGTGAGTACGTGCAAGTCTGTGACTCGGACACA 720
Db 696 ACAGCGCTTCAAGCGCTCGGAGATTCGGTGAGTACGTGCAAGTCTGTGACTCGGACACA 755
Qy 721 AGGTTGGACCCCATGGCACTGCTGGAGCTCGTGGCGGTACTGGACGAGGACCCCGGGTA 780
Db 756 AGGTTGGACCCCATGGCACTGCTGGAGCTCGTGGCGGTACTGGACGAGGACCCCGGGTA 815
Qy 781 GGGGCTGTGTGGGAGCGTGGGATCCTTAAACCTCTGGAGTCTGGGTCGAGCTTCCCTA 840
Db 816 GGGGCTGTGTGGGAGCGTGGGATCCTTAAACCTCTGGAGTCTGGGTCGAGCTTCCCTA 875
Qy 841 AGCAGCGTGCATACTGGGAGCGCTTCAATGTGGAGCGGCTTGTGAGAGTACTTCCAC 900
Db 876 AGCAGCGTGCATACTGGGAGCGCTTCAATGTGGAGCGGCTTGTGAGAGTACTTCCAC 935
Qy 901 TGTGTATCTGTCATCAGCGGCTCTCTAGCGCTTATATAGGAATAAATCTTGTGAGAGT 960
Db 936 TGTGTATCTGTCATCAGCGGCTCTCTAGCGCTTATATAGGAATAAATCTTGTGAGAGT 995
Qy 961 CTGAGGCGCTGGTACAAACAGAGTTCCTTGGGTAACCACTGTATCTTTTGGGGATGACCGG 1020
Db 996 CTGAGGCGCTGGTACAAACAGAGTTCCTTGGGTAACCACTGTATCTTTTGGGGATGACCGG 1055
Qy 1021 CACCTCACCAACCGCATGCTGAGCATGGTATGCTACCAAGTA 1064
Db 1056 CACCTCACCAACCGCATGCTGAGCATGGTATGCTACCAAGTA 1099
```

```
; Sequence 5659, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5659
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-5659
```

Query Match 99.6%; Score 1060.8; DB 4; Length 2087;
Best Local Similarity 99.8%; Pred. No. 4.6e-203;
Matches 1062; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy 1 ATGAGACAGCAGGACGGCCCAAGCCCACTCTCTGAGCGCGCGCTGCTGGGCGCTGGCC 60
Db 36 ATGAGACAGCAGGACGGCCCAAGCCCACTCTCTGAGCGCGCGCTGCTGGGCGCTGGCC 95
Qy 61 CGGAGGGTGTGACCATGCGCTTGGCCCTGCTCATCTCTGGGCTCATGACCTGGGCGCTAC 120
Db 96 CGGAGGGTGTGACCATGCGCTTGGCCCTGCTCATCTCTGGGCTCATGACCTGGGCGCTAC 155
Qy 121 GCCCGCGGGTGGCGCGCTTGGCGTACCGGCTCTTGGGCTCTGGGCTCTTACGGG 180
Db 156 GCCCGCGGGTGGCGCGCTTGGCGTACCGGCTCTTGGGCTCTTACGGG 215
Qy 181 GCCTCTCTTTTACGCGACCTGGTGGCGAGCGCTCTTGGGCTACCTGGAGCACCGGCGG 240
Db 216 GCCTCTCTTTTACGCGACCTGGTGGCGAGCGCTCTTGGGCTACCTGGAGCACCGGCGG 275
Qy 241 GTGGCGGCGGCGCGGGGCGCGCTGGATGCAAGCGCGCGAGTGTGGCGCTGAGC 300
Db 276 GTGGCGGCGGCGCGGGGCGCGCTGGATGCAAGCGCGCGAGTGTGGCGCTGAGC 335
Qy 301 ATCTCCGCTTACAGGAGGACCCCGCGTACCTGGCGCAGTGCCTGGCGTCCGCGCGCC 360
Db 336 ATCTCCGCTTACAGGAGGACCCCGCGTACCTGGCGCAGTGCCTGGCGTCCGCGCGCC 395
Qy 361 CTGCTGTACCGCGCGCGCTGGCGTCTCTCATGTGTGGTGGATGGCAACCGCGCGCAG 420
Db 396 CTGCTGTACCGCGCGCGCTGGCGTCTCTCATGTGTGGTGGATGGCAACCGCGCGCAG 455
Qy 421 GACCTCTACATGCTGCAATGCTTGGCGAGTCTTGGCTGACGAGGACCCCGCACGCTAC 480
Db 456 GACCTCTACATGCTGCAATGCTTGGCGAGTCTTGGCTGACGAGGACCCCGCACGCTAC 515
Qy 481 GTGGGACGGCAACTACCAAGCGCTTGGGAAACCGCGCGGGCGCGGCTGGGCGCC 540
Db 516 GTGGGACGGCAACTACCAAGCGCTTGGGAAACCGCGCGGGCGCGGCTGGGCGCC 575
Qy 541 GGAGCGCTATCGGAGGTGGAGGCGGAGGATCCTTGGGCGGCTGGCGAGTGGAGCGCTGGTG 600
Db 576 GGAGCGCTATCGGAGGTGGAGGCGGAGGATCCTTGGGCGGCTGGCGAGTGGAGCGCTGGTG 635
Qy 601 AGGACTCGAGGTGGTGTGGTGGCGAGCGCTGGGCGGGGCGGAGCGGAGGTCTATGTAC 660
Db 636 AGGACTCGAGGTGGTGTGGTGGCGAGCGCTGGGCGGGGCGGAGCGGAGGTCTATGTAC 695
Qy 661 ACAGCGCTTCAAGCGCTCGGAGATTCGGTGAGTACGTGCAAGTCTGTGACTCGGACACA 720
```


Db 696 ACAGCCTTCAAGCGGCTCGGAGATTGCGTGGGACTAGTGCAGGTCGTGACTCGGACACA 755
Qy 721 AGGTTGGACCCCATGCACTGCTGGAGCTCGTGCGGGTACTGGACGAGGACCCCGGGTA 780
Db 756 AGGTTGGACCCCATGCACTGCTGGAGCTCGTGCGGGTACTGGACGAGGACCCCGGGTA 815
Qy 781 GGGGCTGTTGGTGGGACCTGGGATCCTTAACCTCTGGAATCTCTGGGTGAGTTCCTTA 840
Db 816 GGGGCTGTTGGTGGGATGCGGATCCTTAACCTCTGGAATCTCTGGGTGAGTTCCTTA 875
Qy 841 AGCAGCTGCGATACCTGGGTAGCCTCAATGTGGAGCGGCTTGTCAGAGCTACTTCCAC 900
Db 876 AGCAGCTGCGATACCTGGGTAGCCTCAATGTGGAGCGGCTTGTCAGAGCTACTTCCAC 935
Qy 901 TGTGTATCTGTCATCAGCGGCTCTTAGGCTTATATAGGAATACTCTTGACAGAGTTT 960
Db 936 TGTGTATCTGTCATCAGCGGCTCTTAGGCTTATATAGGAATACTCTTGACAGAGTTT 995
Qy 961 CTTGAGGCTGTTACAAACAGAGTTCCTGGGTACCCACTGTTTGGGGATGACCGG 1020
Db 996 CTTGAGGCTGTTACAAACAGAGTTCCTGGGTACCCACTGTTTGGGGATGACCGG 1055
Qy 1021 CACCTCACCAACCGCATGCTCAGCATGGGTTATGTACCAAGTA 1064
Db 1056 CACCTCACCAACCGCATGCTCAGCATGGGTTATGTACCAAGTA 1099

RESULT 3

US-09-155-768-1
; Sequence 1, Application US/09155768A
; Patent No. 6162908
; GENERAL INFORMATION:
; APPLICANT: SEIKAGAKU CORPORATION
; TITLE OF INVENTION: POLYPEPTIDE OF HYALURONAN SYNTHASE OF
; TITLE OF INVENTION: HUMAN-ORIGIN AND DNA ENCODING THE SAME
; FILE REFERENCE: TOYAM37.001APC
; CURRENT APPLICATION NUMBER: US/09/155,768A
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: JP 8-084326
; EARLIER FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: JP 8-109663
; EARLIER FILING DATE: 1996-04-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2117
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (149)...(1777)
US-09-155-768-1

Query Match 98.7%; Score 1051.6; DB 3; Length 2117;

Best Local Similarity 99.6%; Pred. No. 3.2e-201;

Matches 1054; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 7 CAGCAGGACGCGCCCAAGCCACTCTGCAGCCGCGCTGCTCGGCTGCGCGGAGG 66
Db 50 CAGCAGGACGCGCCCAAGCCACTCTGCAGCCGCGCTGCTCGGCTGCGCGGAGG 109
Qy 67 GTGCTGACCATCGCCTTGGCCCTGCTCATCTCGGCGCTCATGACCTGGGCTTACGCCGCC 126
Db 110 GTGCTGACCATCGCCTTGGCCCTGCTCATCTCGGCGCTCATGACCTGGGCTTACGCCGCC 169
Qy 127 GGGGTGCGGCTGGCTCCGATCGCTACGCGCTTCTGGCGCTTACGCGGCTTACGCGGCTTC 186
Db 170 GGGGTGCGGCTGGCTCCGATCGCTACGCGCTTCTGGCGCTTACGCGGCTTACGCGGCTTC 229
Qy 187 CTTTCAGCGCACCTGGTGGCGAGAGCTCTTCGGGTACCTTGAGACCGCGGCTGGG 246
Db 230 CTTTCAGCGCACCTGGTGGCGAGAGCTCTTCGGGTACCTTGAGACCGCGGCTGGG 289

RESULT 4

US-09-949-016-17401
; Sequence 17401, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

Qy 247 GCGCGCGCGCGCGGCGCTGATGACACCAACCGCGCGAGTGTGGCGCTGACCACTCTCC 306
Db 290 GCGCGCGCGCGCGGCGCGCTGATGACACCAACCGCGCGAGTGTGGCGCTGACCACTCTCC 349
Qy 307 GCCTACACGAGGAGGACCCCGCGTACCTGGCGCAGTGCCTGGCGTCCGCGCGCGCGCTGCTG 366
Db 350 GCCTACACGAGGAGGACCCCGCGTACCTGGCGCAGTGCCTGGCGTCCGCGCGCGCGCTGCTG 409
Qy 367 TACC CGCGCGCGCGCTGCGGTCTCTCATGTTGGTGGATGGCAACCGCGCGCGAGACCTTC 426
Db 410 TACC CGCGCGCGCGCGCTGCGGTCTCTCATGTTGGTGGATGGCAACCGCGCGCGAGACCTTC 469
Qy 427 TACATGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCGCGCGAGTGTGG 486
Db 470 TACATGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCGCGCGAGTGTGG 529
Qy 487 GACGCAACTACCAACGAGCCTTGGGAAACCCCGCGCGCGCGCGCGGTGGCGCGCGAGCC 546
Db 530 GACGCAACTACCAACGAGCCTTGGGAAACCCCGCGCGCGCGCGCGGTGGCGCGCGAGCC 589
Qy 547 TATCGGAGGTGGAGGCGGAGATCCTGGCGGCTGGGAGTGGAGGCGCTGCTGTAGAGT 606
Db 590 TATCGGAGGTGGAGGCGGAGATCCTGGCGGCTGGGAGTGGAGGCGCTGCTGTAGAGT 649
Qy 607 CCGAGGTGCGTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCTGTACACGCC 666
Db 650 CCGAGGTGCGTGTGCGTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCTGTACACGCC 709
Qy 667 TTCAAGGCGCTCGAGAGATTCGGTGGACTACGTCGAGGTCTGTGACTCGGACACCAAGTTG 726
Db 710 TTCAAGGCGCTCGAGAGATTCGGTGGACTACGTCGAGGTCTGTGACTCGGACACCAAGTTG 769
Qy 727 GACCCCATGCGCACTGCTCGAGCTCGTGGGCTGCTGGAGCAGGAGCCCGCGGTAGGGCT 786
Db 770 GACCCCATGCGCACTGCTCGAGCTCGTGGGCTGCTGGAGCAGGAGCCCGCGGTAGGGCT 829
Qy 787 GTTGGTGGGACGCTCGGATCCTTAACCCCTCTGGAATCTCTGGGTGAGTCTTCAAGCAGC 846
Db 830 GTTGGTGGGATGTCGGATCCTTAACCCCTCTGGAATCTCTGGGTGAGTCTTCAAGCAGC 889
Qy 847 CTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGCTTGTGAGACTCTTCCACTGTGTA 906
Db 890 CTGCGATACCTGGGTAGCCTTCAATGTGGAGCGGCTTGTGAGACTCTTCCACTGTGTA 949
Qy 907 TCCTGCATCAGCGGTCTCTAGGCTATATAGGAATAACCTCTTGCAGCAGTTCTTTCAG 966
Db 950 TCCTGCATCAGCGGTCTCTAGGCTATATAGGAATAACCTCTTGCAGCAGTTCTTTCAG 1009
Qy 967 GCCTGGTACAAACGAGATTCCTGGGTACCCACTGTACTTTTGGGGATGACCGGACCTC 1026
Db 1010 GCCTGGTACAAACGAGATTCCTGGGTACCCACTGTACTTTTGGGGATGACCGGACCTC 1069
Qy 1027 ACCAACCAGATGCTCAGCATGGGTTATGCTACCAAGTA 1064
Db 1070 ACCAACCAGATGCTCAGCATGGGTTATGCTACCAAGTA 1107

	Best Local Similarity	62.9%; Pred. No. 1.le-61;	Matches 633; Conservative 0; Mismatches 335; Indels 39; Gaps 4;
Qy	61	CGAGGGTGTGACCATCGCTTCGCCCTTGCTCATCTCTGGGCCTCATATCACTCTGGCCCTAC	120
Db	28	CGTGTGTGGCACCAAGCCTTTTGCCCTGGCAGTGTGGGTGGCATCTCTGGCAGCCTAT	87
Qy	121	GCCGCCGGGGTGCCTGGCTTCGATCGCTTAAGCCCTCTCTGGCCTTCGGCCTTCACGGG	180
Db	88	GTACGGGCTACCAGTTTCATCCACACGGAAGAAGCACTACCTGTCTCTTCGGCTGTACGGC	147
Qy	181	GCCTTCCTTTTCAGCGCAGCTGTGGCGACAGAGCCTCTTCGCTACCTGGAGACACCGCGG	240
Db	148	GCAATCTCTGGGCTTGCAACTGCTCATTCAGAGCCTTTTTCCTTCTCTGGAGCACCGGCGC	207
Qy	241	GTGGCGGCGCGCGCGGGGCGCTGGATTGCAGCCACCGCGC---GCAGTGTGGCGCTG	297
Db	208	ATGCGACGTGCCGCGACAGCCCTGAAGCTGCCCTCCC CGCGGGGCGCTCGTGGCACTG	267
Qy	298	ACCATCTCCGCTACACAGGAGGACCCCGCGTAACTCTGCCAGTGTCTGGGGTTCGCCCGC	357
Db	268	TGCAATTGCCGCATAACAGGAGGACCTTGACTACTTTCGCAAGTGTCTCTGGCGCCAG	327
Qy	358	GCCTGTCTTACCGCGCGCGCGCTGCGGCTCTCATGTGTGGTGAATGCMAACCGCGCC	417
Db	328	CGCATCTCTTCCCTGAC-----CTCAAGGTGTGCATGTGTGGTGAATGCMAACCGCGC	381
Qy	418	GAGGACCTTACATGTGTGCATATGTTCCGGCAGGTCTTCGTCTACAGAGAACCCCGCCACG	477
Db	382	GAGGACGCTACATGCTGGACATCTTCCAAGAGTGTCTGGCGGACCCGAGCAGCGCGC	441
Qy	478	TACGTGTGGACCGGCAACTACCAACAGCCTTGGGAACCCCGCGCGCGCGGTGGCGG	537
Db	442	TTCTTTGTGTGGCGCAGCAACTTCCA-----TGAGGCGAGCGAGGGTGAG	486
Qy	538	GCCGGAGCCTATCGGAGGTGGAGCGGAGGATCCTGGCGGCTGGGCAGTGGAGGCGCTG	597
Db	487	ACGAGGCCAGCCTGCAAGAGGGGATGGACCGT-----GTGCGGGATGTG	531
Qy	598	GTGAGGACTCGCAGGTGCTGTGGCGCAGCGCTGGGCGCGCAAGCGCAGGTCAATG	657
Db	532	GTGGCGGCAGACCTTCTCTGTGATCATGCAGAAGTGGGAGGCGAAGCGCGAGGTCAATG	591
Qy	658	TACACAGCCTTCAAGGCGCTCGGAGATTCCGTGTGACTACGTGCAGTGTCTGTGACTCGGAC	717
Db	592	TACACGCGCTTCAAGGCGCTCGGCGATTCCGTGTGACTACATCCAGGTGTGCGACTCTGAC	651
Qy	718	ACAAGTTTGACCCCATGGCACTGCTCGAGCTCGTGGGGTACTTGGAGGAGGACCCCCGG	777
Db	652	ACTGTCTGTGATCCAGCTCGACCATCGAGATGCTTCGAGTCTCTGGAGGAGGATCCCAA	711
Qy	778	GTAGGGCTGTGTGTGGGAGCTGCGGATCCTTAACCTCTCGACTCTCGGTCAGCTTC	837
Db	712	GTAGGGGGAGTGGGGGAGATGTCCAGATCCTCAACAAGTACGACTCATGGATTTCTTTC	771
Qy	838	CTAAGCAGCTTCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTCTAGAGCTACTTC	897
Db	772	CTGAGCAGCGTGCGTACTGGATGGCCTTCAACGTGGAGCGGGCTGCCAGTCTCTACTTT	831
Qy	898	CATGTGTATCTTCATCAGGGTCTCTTAGGCTATATAGGATATACCTCTTTCGAGCAG	957
Db	832	GGCTGTGTGAGTGTATTTAGTGGGCGCTTGGGCATGTACCGCAACAGCCTCTCCACGAG	891
Qy	958	TTTCTTGAGCGCTGGTACAAACAGAAAGTCTCTGGGTACCCACTGTACTTTTTGGGATGAC	1017
Db	892	TTCTTGAGGACTGGTACCATCATAGAAAGTCTTAGGAGCAAGTGACGTCTCGGGGATGAC	951
Qy	1018	CGGCACTTCCAAACCGCATCTCTCAGCATCGGGTTATGCTACCAAGTA	1064
Db	952	CGGCACTTCCAAACCGAGTCTGTAGCCTTGGCTTACCGAATCTAAGTA	998

RESULT 7

US-08-812-008-31
; Sequence 31, Application US/08812008
; Patent No. 6602693
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustine, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: U.S.A.
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,008
; FILING DATE: 05-MAR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/675,499
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Embretson, Janet E
; REGISTRATION NUMBER: 39,665
; REFERENCE/DOCKET NUMBER: 150.183US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: cdNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...1662
; OTHER INFORMATION:
US-08-812-008-31

Query Match	29.3%	Score 311.6;	DB 4;	Length 1665;
Best Local Similarity	60.3%;	Pred. No. 2.1e-53;		
Matches 609;	Conservative 0;	Mismatches 359;	Indels 42;	Gaps 4;
Qy	61	CGAGGGTGCTGACATCGCCTTCGCCCTGCTCATCTCGGCGCTCATGACCTGGCGCCTAC	120	
Db	28	CGTGTGTGGGCACCAAGTCTGTGTTCCTCGTAGTGTGGGAGGATCCTGGCGCCTAT	87	
Qy	121	GCOCGCGGGGTGCGCTGGCGCTCGATCGCTACGGCCCTCTCGGCGCTTCGGCGCTACGGG	180	
Db	88	GTGACAGGCTACAGTTTATCCACACAGAAAGCACTACCTGTCTCTTGGCGCTCTACGGT	147	
Qy	181	GCCTTCCTTTTCAGCGACCTGTGTGGCGACAGACCTCTTCGAGTACCTGGAGCACCGCGG	240	
Db	148	GCCATCCTGGGTCTACATCTGCTATCCAGAGCCTGTTTGCTTCTTCGGAGCACCGTGA	207	
Qy	241	GTGGCGCGCGCGCGCG-----GGGGCCGCTGGATGCAGCCACCGCGCGAGTGTGGCG	294	
Db	208	ATGGCAGGCGAGGGCGCCCCCTCACTGCATGCTCCAGAGGTGCGGTTCAGTGCCA	267	
Qy	295	CTGACCATCTCCGCCTACCAAGGAGGACCCCGCGTACCTGCGGCCAGTGCCTGGCGTCGCC	354	
Db	268	CTCTGCACTTCCTGCTACCAAGAGGACCCCGAATCTCTGCGCAGTGCCTTCGCTCAGCT	327	

```
QY 355 CGCGCCCTGCTACCGCGCGCGCTGCGCTCCTCATGGTGGTGGCAACGC 414
Db 358 CAGCGCATGGCTTTTCCAAC-----CTCAAGGTGGTCATGGTAGTGGCAATCGC 381
QY 415 GCGGAGGACCTTACATGGTGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCC 474
Db 382 CAGGAAGATACCTACATGTTGACATCTTCATGAGGTGCTGGGTGGCACTGAGCAAGCT 441
QY 475 AGCTACGTGTGGAGCGCAACTACCAACGAGCCCTGGGAACCGCGCGCGCGCGCGGTG 534
Db 442 GCTTCTTTTGTGGCGTAGCAATTTCCA-----TGAGCGCGGTGAAGGA 486
QY 535 GCGCGCGAGCTACTCGGAGGTGAGCGCGGAGGATCTTGGCGGCTGGCAGTGGAGCG 594
Db 487 GAGACAGAGCGAGCTCGGAGGAGGATGGAGCGT-----GTGCGAGCT 531
QY 595 CTGCTGAGGACTCGCAGGTGCGTGTGCTGGCGCAGCGCTGGGCGCGCAAGCGCGAGTGC 654
Db 532 GTGGTGTGGCGCAGCACTTCTCATGTCATGTCAGAAAGTGGGGGGCAAGGTGAGTGC 591
QY 655 ATGTACAGGCTTCAAGCGCTCGGAGATTGGGTGGAATCGGTGCAAGTCTGTGACTCG 714
Db 592 ATGTACAGCTGCTTCAAGCGCTTGGCAACTCAGTGGACTACATCCAGGTGTGTGACTCT 651
QY 715 GACACAGGTGGACCCCATGGCACTGCTGGAGCTCGTGGGTACTGGACGAGACCCC 774
Db 652 GACACTGTGCTGGACCCCATGGCACTGACCATTTGAGATGCTTGGAGTCTTGGAAAGAGATCCC 711
QY 775 CGGTAGGGGTGTTGGTGGGAGCTGGGAGCTGCGGATCTTAAACCTCTGGAATCTCTGGGTGAGC 834
Db 712 CAAGTAGAGGTGTTGGAGGAGATGTCAAATCTTCAACAAGTATGATTCATGATCTCC 771
QY 835 TTCTTAAGCAGCTCGATACCTGAGTACCTTCAATGTGGAGCGGGCTTGTCAAGTAC 894
Db 772 TTCTTGAAGGTGAGGTGAGTACCTGAGTGGCTTCAACGTGGAGCGGGCTGCGAGTCTAC 831
QY 895 TTCCACTGTATCTGTCATCAGCGCTCTAGGCTTATAGCCCTATATAGAAATAAATCTTTCAG 954
Db 832 TTGGCTGTGTCAATGATGATGAGTGGCTTTGGGCACTGACCGCAACAGGCTTCTTCAG 891
QY 955 CAGTTCTTGGGCTGCTGACCAACAGAAAGTTCTTGGGTACCCACTGTACTTTTGGGGAT 1014
Db 892 CAGTTCTTGGAGGATTTGATCAATCAGAAAGTTCTTAGGCAAGTGCAGCAAGTGTGGGAT 951
QY 1015 GACCGCACTTCAACCAACGATGCTCAGCATGAGTGGGTATGTTACCAAGTA 1064
Db 952 GATCGGCACCTTACCAACCGAGTCTGAGTCTTGGCTACCGGACTTAAGTA 1001
```

RESULT 8

```
US-08-675-499A-1
; Sequence 1, Application US/08675499A
; Patent No. 6492150
```

GENERAL INFORMATION:

```
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESS: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
```

COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2947 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-675-499A-1
```

Query Match 23.2%; Score 246.8; DB 4; Length 2947;

Best Local Similarity 57.8%; Pred. No. 2e-40;

Matches 524; Conservative 0; Mismatches 337; Indels 45; Gaps 3;

```
QY 159 CCTGGCCCTTCAGCGGCTTCTTCTTACGCGACCTGTGTGGCGCAGAGCCTCTT 218
Db 636 CTTCTCATTTGGACTGTAGCGTGCCTTTTAGCTCGCATCTCATCATCCAAAGCCTCTT 695
QY 219 CGGTACTCTGGACACCGCGGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 278
Db 696 TGCCTTTTGGAAACACCGGAAATGAAGAAATCCCTTTGAAAACCCCGATTAAATGAACAA 755
QY 279 CGCGCGCAGTGTGGCGTGCACCATCTCCGCTACAGAGGAGACCCGCGTACCTGCGGCA 338
Db 756 AAG-----GTAGCATCTGTCATCGTGTGATCAACAGAGACCTGACTACTTACGAA 809
QY 339 GTGCTCGGCTGCGCCCGCGCTGTGTATCCCGCGCGCGCGCTGTGCGCTCTCATGGT 398
Db 810 ATGTTTGAATCTGTGAAAGGCTGACTACCTCTG-----GGATTAAAGTCTGTGATGGT 863
QY 399 GGTGATGGCAACCGCGCGGAGACCTCTACATGTCGACATGTCGCGAGGTCTTCGC 458
Db 864 CATCGATGGGAACCTCAGACGACCTTTACATATGACATATTTACGCGAAGTTATGG 923
QY 459 TGACGAGGACCCCGCCACGTACGTGTGGAGCGGCAACTACCACCGCGCTGGGAACCCGC 518
Db 924 CAGGACAAATCGGCCACGTACATCTGGAAGAACAACTTTTCATGAAAG----- 972
QY 519 GCGCGCGCGCGGTGGCGCGCGAGCCTATCGGGAGGTGGAGCGGAGGATCTCTGGCG 578
Db 973 -----GGACCTGGTGACAGAGAGTCCCATAAAGAAAGTTC 1010
QY 579 GCTGCGATGGAGCGCTGTGTGAGGACTCGCAGAGTGTGTGCGTGGCGCAGCGCTGGG 638
Db 1011 ACAACATGTCAACCAATTGGTCTTTGTCTAAACAAAGTATTTGTCATCATGCAAAAATGGG 1070
QY 639 CGCAAGCGGAGGTGATGACAGCGCTTCAAGCGCTCGGAGATTCCGTTGGACTACGT 698
Db 1071 TGAAGAGAGAGAGTCAATGATACAGCGCTTCAAGAGCACTGGGGGGAAGCGTGGATTATGT 1130
QY 699 GCAGGTCTGTGACTCGGACCAAGGTTGGACCCCATGCGCACTGTGTGAGTCTGTGCGGGT 758
Db 1131 ACAGTGTGTGACTCAGATCTATGCTTGACCTGCTCTCTCTGTGGAGATGGTGAAGGT 1190
QY 759 ACTGACGAGGACCCCGGTAGGGCTGTGTGGGAGCGTGGCGGATCTTAAACCTCT 818
Db 1191 CTTAGAGGAAGACCTTATGGTGGAGGTGTGGAGGAGATGTCCAGATTTTAAACAAAGTA 1250
QY 819 GGAATCTCTGGGTGAGTCTTCTTAAGCAGCTGCGGATCTGCGGTAGCTTCAATGTGGAGCG 878
```



```

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201780
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-201780

```

```

Query Match 21.6%; Score 230.4; DB 4; Length 601;
Best Local Similarity 98.7%; Pred. No. 3.2e-37;
Matches 231; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 696 CGTGCAGCTCTGTGACTCGGACACAGAGTTGGACCCGATGGCACTGCTGGAGCTCGTGG 755
DB 403 CTTACAGCTCTGTGACTCGGACACAGAGTTGGACCCGATGGCACTGCTGGAGCTCGTGG 344
QY 756 GGTACTCGACAGAGACCCCGGGTGGGCTGTGGTGGGACGCTGGGATCCTTAACCC 815
DB 343 GGTACTCGACAGAGACCCCGGGTGGGCTGTGGTGGGACGCTGGGATCCTTAACCC 284
QY 816 TCTGACCTCTGGGTCAGCTTCTTAAGCAGCTGCGATCTGCTGAGTGGTCAATGTGGA 875
DB 283 TCTGACCTCTGGGTCAGCTTCTTAAGCAGCTGCGATCTGCTGAGTGGTCAATGTGGA 224
QY 876 GGGGCTTGTGACAGCTACTTCCACTGTATCTGCTGATCAGCGCTGCTTAGG 929
DB 223 GGGGCTTGTGACAGCTACTTCCACTGTATCTGCTGATCAGCGCTGCTTAGG 170

```

```

RESULT 11
US-09-949-016-737
; Sequence 737, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 737
; LENGTH: 3003
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-737

```

```

Query Match 20.8%; Score 221; DB 4; Length 3003;
Best Local Similarity 56.2%; Pred. No. 2.9e-35;
Matches 506; Conservative 0; Mismatches 350; Indels 45; Gaps 3;

QY 164 CTTTCGGGCTCTACGGGCTCTTCTTTCAGCGCACTGCTGGCGCAGAGCTCTTGGCGT 223
DB 669 CTTTGGACTGTATGGTGGCTTTTGGCATCACACCTCATATCCAAAGCTGTTGGCT 728
QY 224 ACCTGAGACACGGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 283
DB 729 TTTTGGAGCAGCGAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 782

```

```

QY 284 GCAGTGTGGCGCTGACCATCTCCGCTACACAGAGGACCCGCGCTACCTGGCCAGTGCC 343
DB 783 AAACAGTTGGCCCTTTGCATCGCTGCTATCAAGAAGATCCAGACTACTTAAAGAAATGT 842
QY 344 TGGCGTCCGCGCGCCCTGCTGTACCGCGCGCGCGCTGCGGCTCTCATGTGGTGG 403
DB 843 TGCAATCTGTGAAAGGCTAACCTACC-----CTGGGATTAAGATTTGTCATGGTCAATG 896
QY 404 ATGCAACCGCGCGAGGACCTCTACATGCTGCAATGTTCCGGGAGGCTTCCGCTGACG 463
DB 897 ATGGGAATCTCAGAAAGATGACCTTTACATGATGGACATCTTCAGTGAAGTATGAGGCA 956
QY 464 AGGACCCCGCACGCTAGCTGTGGAGCGCACTACCAACGACCTTCCAGGAAAG----- 523
DB 957 ACAATCAGCCACTTATATCTGGAGAACAACTTCCAGGAAAG----- 1000
QY 524 CGGCGCGGTGGCGCGGAGCTTATCGGAGGTGGAGCGGAGGATCCTGGGCGGTGG 583
DB 1001 -----GGTCCCGGTGAGACAGATGAGTACATAAAGAAAGCTCGCAAC 1043
QY 584 CAGTGGAGGCGCTGCTGAGGACTCGCAGGTGCTGTGGTGGCGCAGCGCTGGGCGGCA 643
DB 1044 ACGTAAACGCAATGGTCTTGTCCAAACAAAGTATCTGCATCATGCAAAATGGGCTGAA 1103
QY 644 AGCGGAGGTCATCTACACAGCTTCAAGGCGCTCGAGATTCGCTGCACTACGTGACG 703
DB 1104 AAAGAGAAGTCAATGTACACAGCTTACAGCACTGGGACGAGTGTGGATTTATGTACAG 1163
QY 704 TCTGTGACTCGGACACAAAGTTGGACCCCACTGGCACTGCTGGAGCTCGTGGGCTAC 763
DB 1164 TTTGTGATTCAGACACTATGCTTGACCCAGCTCATCTGTGGAGATGTTAAAGTATTTAG 1223
QY 764 ACAGAGACCCCGGTAGGCGCTGTTGGTGGGAGCTGGGATCCTTAACCTCTGGACT 823
DB 1224 AAGAAGATCCCATGTTGGAGGTGTTGGGGAGATGTCCAGATTTTAAACAAAGTACGATT 1283
QY 824 CTTGGGTGAGCTTCTTAAGCAGCTGCGATACTGGGTAGCTTCAATGTGGAGCGGCT 883
DB 1284 CTTGATCTATCTCAGCAGTGTAGATATTGGATGGCTTTTAAATATAGAAAGGCT 1343
QY 884 GTCAGAGCTACTTCCACTGTGTATCTGCACTACGCGTCTCTAGGCTCTATAGGAATA 943
DB 1344 GTCAGTCTTATTTTGGGTGTGTTTCAAGTGCATTTAGTGGACCTCTGGGAATGTACAGAA 1403
QY 944 ACCTTTCGACAGCTTTCTTGGAGCTGTTACACAGCAAGTCTCTGGGTACCACTGTA 1003
DB 1404 CTTTGTGTCATGATGTTTGTGGAAGATTGGTACAAATCAAGAAATTTATGGGCAACCAAT 1463
QY 1004 CTTTGGGATGACCGGCACTCACCACCGCATGCTCAGCATGGGTATGTACCAAGT 1063
DB 1464 GCTTGTGATGACAGGATCTCAGAACCGGCTGCTGAGCTGGGCTATGCACAAAT 1523
QY 1064 A 1064
DB 1524 A 1524

```

```

RESULT 12
US-09-949-016-15470
; Sequence 15470, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```


; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15470
; LENGTH: 9871
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15470

Query Match 15.7%; Score 167.6; DB 4; Length 9871;
Best Local Similarity 58.9%; Pred. No. 1.5e-24;
Matches 384; Conservative 0; Mismatches 22; Indels 39; Gaps 4;
QY 61 CGAGGGTGTGACCATCGCCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
DB 2028 CGTGTGTGGGACACGAGCTGTTGGCTTGGCAGTGTGGTGGCATCTGGCAGCCTAT 2087
QY 121 GCCCGGGGTGCGCTGGCTCCGATCGCTAGGCGCTCTGGCCTTGGGCTTACGGG 180
DB 2088 GTGACGGGCTACCAAGTTTCATCCACACGGAAAGCACTACCTGTCTTGGGCTGTACGGC 2147
QY 181 GCCTTCTTTTTCAGCGCACCTGTGTGGCGAGAGCCTCTTGGCTTCTGAGCAGCGCGG 240
DB 2148 GCCATCTTGGGCTGACCTGCTCATTCAGAGCCTTTTGGCTTCTGAGCAGCGCGC 2207
QY 241 GTGGCGGGCGGGCGGGCGGCTGGATGAGCCACCGCGC---GCAGTGTGGCGCTG 297
DB 2208 ATGGAGCTGCGGGCAGGCCCTGAAGCTGCCCTCCCGGGGGGCTCGTGGCACTG 2267
QY 298 ACCATCTCGGCTTACAGAGAGACCCCGGTACTGTGGCGCATGTGCTGGGCTCGCGCGC 357
DB 2268 TGCATTGCGCATACAGAGAGACCTGACTACTTGGCAAGTGTGCTGCTCGGCTCGGCCAG 2327
QY 358 GCCTGTGTACCGCGCGCGGGCTGCGCTCTCATGTGTGGATGGCAACCGCGCC 417
DB 2328 CGCATCTCTCTCCCTGAC-----CTCAAGTGTGTGGTGGATGGCAACCGCGAG 2381
QY 418 GAGGACCTTACATGCTGACATGTTCCGCGAGGTCTTCCGCTGAGAGAGACCCCGCCACG 477
DB 2382 GAGGAGCGCTACATGCTGACATCTTCCAGAGGTGCTGGCGGACCGAGCAGCGCGC 2441
QY 478 TACGTGTGGAGCGGAACTACCAACAGCCCTGGGAAACCCCGCGCGGGCGGCGGTGGGC 537
DB 2442 TTCTTTGTGTGGCGAGCACTTCAT-----GAGGCGAGCGAGGTGAG 2486
QY 538 GCGGAGCTATCGGAGGTGGAGCGGAGGATCTTGGCGGCTGGCAGTGGAGCGCTG 597
DB 2487 ACGGAGCGAGCTTCAGAGGGCATGACCGT-----GTGGGATGTG 2531
QY 598 GTGAGGACTCGCAGGTGCTGTGCTGGCGAGCGCTGGGCGGCAAGCGAGGTCTATG 657
DB 2532 GTGGGGCCAGCACCTTCTGTGCATCATGCAAGTGGGAGGCAAGCGGAGGTCTATG 2591
QY 658 TACAGAGCTTCAAGCGCTGGAGATTCGTGGACTACGTGCAAGTCTGTG 709
DB 2592 TACAGGCTTCAAGGCGCTGGCGATTCTGGTGGACTACATCCAGGTAAGG 2643

RESULT 13
US-09-949-016-132199
; Sequence 132199, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132199
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-132199

Query Match 12.9%; Score 137.4; DB 4; Length 601;
Best Local Similarity 58.5%; Pred. No. 1.2e-18;
Matches 300; Conservative 0; Mismatches 201; Indels 12; Gaps 3;
QY 61 CGGAGGGTGTGACCATCGCCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
DB 50 CGTGTGTGGGACACGAGCTGTTGGCTTGGCAGTGTGGTGGCATCTGGCAGCCTAT 109
QY 121 GCCCGGGGTGCGCTGGCTCCGATCGCTACGCGCTCTTGGGCTTCTGGGCTTCTACGGG 180
DB 110 GTGACGGGCTACCAAGTTTCATCCACACGGAAAGCACTACCTGTCTTGGGCTGTACGGC 169
QY 181 GCCTTCTTTTTCAGCGCACCTGTGTGGCGAGAGCCTCTTGGCTTCTGAGCAGCGCGG 240
DB 170 GCCATCTTGGGCTGCACTGCTCATTCAGAGCCTTTTGGCTTCTTGGAGCAACCGCGC 229
QY 241 GTGGCGGGCGGGCGGGCGGCTGGATGAGCCACCGCGC---GCAGTGTGGCGCTG 297
DB 230 ATGGAGCTGCGGGCAGGCCCTGAAGCTGCCCTCCCGCGGGGGCTCGTGGCACTG 289
QY 298 ACCATCTCGGCTTACAGAGAGACCCCGGTACTGTGGCGCATGTGCTGGGCTCGCGCGC 357
DB 290 TGCATTGCGCTACAGAGAGACCTGACTACTTGGCAAGTGTGCTGCTCGGCTCGGCCAG 349
QY 358 GCCTGTGTACCGCGCGGGCTGCGCTCTCATGTGTGGATGGCAACCGCGCC 417
DB 350 CGCATCTCTCTCCCTGAC-----CTCAAGTGTGTGGTGGATGGCAACCGCGAG 403
QY 418 GAGGACCTTACATGCTGACATGTTCCGCGAGGTCTT---CGCTGAGAGAGACCCCGCC 474
DB 404 GAGAGCGCTACATGCTGGACATCTTCCAGAGGTGCTGGGCGGACCGAGCAGCGCGC 463
QY 475 ACGTACGTGTGGGACGGCACTACCAACAGCCCTGGGAAACCCCGCGGGCGGCGCGGTG 534
DB 464 TTCTTTGTGTGGCGAGCACTTCATCAGGACGAGGAGGTGAGACGAGGCGAGCCTG 523
QY 535 GCGCGGAGCTATCGGAGGTGGAGCGGAG 567
DB 524 CAGGAGGCGCATGGACCGTGTGGGATGTGGT 556

RESULT 14
US-09-949-016-12479
; Sequence 12479, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12479
; LENGTH: 32176
; TYPE: DNA


```
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(32176)
; OTHER INFORMATION: n = A,T,C or G
US-08-675-499A-25

Query Match      12.0%; Score 127.8; DB 4; Length 32176;
Best Local Similarity 67.4%; Pred. No. 1.6e-16;
Matches 180; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 798 CGTGGGATCCTTAACCCCTCTGGACTCCTGGGTGAGCTTCTTAAGCAGCCCTGCCATCTG 857
Db 27685 CTTGAGATTTTAAACAAGTACGATTCCTGGATCTCATCTCTCAGCAGTGTAAAGATATG 27744

Qy 858 GGTAGCCTTCAATGTGGAGCGGGCTGTGTCAGAGCTACTTCCACTGTGTATCCTGCATCAG 917
Db 27745 GATGGCTTTTAAATATAGAAAGGGCTGTGAGTCTTTATTTTGGGTGTGTTTCAGTGCATTAG 27804

Qy 918 CGGTCTCTAGCCCTATATAGGATTAACCTCTTGCAGCAGTTTCTTGAGGCTGTGTACAA 977
Db 27805 TGGACCTCTGGGAATGTACAGAACTCCTTGTGTGATGAGTTTGTGGAAGATTTGGTACAA 27864

Qy 978 CCAGAAGTTCCTGGGTACCCACTGTACTTTTGGGGATGACGGGCACCTTCACCAACCGCAT 1037
Db 27865 TCAAGAATTTATGGCAACCAATGTAGCTTTTGGTATGACAGGCATCTCAGNACCGGT 27924

Qy 1038 GCTCAGCATGGTTATGTACCAAGTA 1064
Db 27925 GCTGAGCCTGGGCTATGCAACAAATA 27951

RESULT 15
US-08-675-499A-25
; Sequence 25, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 base pairs
```

```
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
US-08-675-499A-25

Query Match      9.0%; Score 95.6; DB 4; Length 235;
Best Local Similarity 71.8%; Pred. No. 2.4e-10;
Matches 125; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 891 CTACTTCCACTGTGTATCTCTGCATCAGCGGTCTCTAGGCCTATATAGGAATAACCTCTT 950
Db 4 CTACTTTGGCTGTGTGCAGTGTATTAGTGGCCCTTGGGCATGTACCGCAACAGCCTCCT 63

Qy 951 GCACAGTTTCTTGAGGCTGTGTACACCAAGAGTTCTCTGGGTACCCACTGTACTTTGG 1010
Db 64 CCAGCAGTTCTTGAGGACTGTGTACCATCAGAAAGTTCTTAGGCAGCAAGTGCAGCTTCGG 123

Qy 1011 GGATGACCGGCACCTCACCACCGCATGCTCAGCATGGGTTATGCTACCAAGTA 1064
Db 124 GGATGACCGGCACCTCACCACCGCATGCTCAGCATGGGTTATGCTACCAAGTA 177

Search completed: March 13, 2005, 08:12:22
Job time : 212.778 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 08:01:42 ; Search time 749.676 Seconds
(without alignments)
8450.560 Million cell updates/sec

Title: US-10-672-399-7

Perfect score: 1065

Sequence: 1 atgagacagcagcgccgcccc.....tgggtatgtacctaacgtaa 1065

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:**

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1065	100.0	1065	18	US-10-672-399-7
2	1064	99.9	1737	18	US-10-672-399-1
3	1064	99.9	2116	13	US-10-042-523-1
4	928	87.1	1083	18	US-10-672-399-5
5	927.4	87.1	1071	18	US-10-672-399-3
6	786.6	73.9	1752	10	US-09-902-939-1
7	422	39.6	662	18	US-10-363-345A-13975
8	422	39.6	662	18	US-10-363-345A-13976
9	421	39.5	490	10	US-09-918-995-27210
c	393.8	37.0	662	18	US-10-363-345A-13973
11	393.8	37.0	662	18	US-10-363-345A-13974

12	353	33.1	1662	17	US-10-295-027-369	Sequence 369, App
13	353	33.1	1662	17	US-10-188-832-136	Sequence 136, App
14	323.8	30.4	1767	17	US-10-309-560-9	Sequence 9, Appli
15	311.6	29.3	1665	10	US-09-902-939-3	Sequence 3, Appli
16	246.8	23.2	1659	10	US-09-902-939-2	Sequence 2, Appli
17	246.8	23.2	4194	15	US-10-262-526-3	Sequence 3, Appli
18	221	20.8	3003	10	US-09-918-624B-64	Sequence 64, Appli
19	221	20.8	3003	15	US-10-262-526-1	Sequence 1, Appli
20	221	20.8	4018	10	US-09-814-353-19942	Sequence 19942, A
21	212.4	19.9	3387	15	US-10-084-817-184	Sequence 184, App
22	212.4	19.9	3387	16	US-10-252-157-84	Sequence 84, Appli
23	100.2	9.4	793	9	US-09-910-943-552	Sequence 552, App
24	95.6	9.0	570	9	US-09-880-107-3449	Sequence 3449, Ap
25	72	6.8	15559	18	US-10-646-664-1	Sequence 1, Appli
26	71	6.7	2182	18	US-10-437-963-62269	Sequence 62269, A
c	70	6.6	43058	9	US-09-954-456-292	Sequence 292, App
c	70	6.6	43058	9	US-09-954-456-329	Sequence 529, App
c	70	6.6	43058	9	US-09-880-107-3950	Sequence 3950, Ap
30	69.4	6.5	5784	14	US-10-152-886-64	Sequence 64, Appli
31	68.6	6.4	1755	18	US-10-437-963-78301	Sequence 78301, A
32	68.4	6.4	2214	15	US-10-156-761-2510	Sequence 2510, Ap
c	68.4	6.4	9025608	15	US-10-156-761-1	Sequence 1, Appli
34	67.6	6.3	1380	18	US-10-425-115-170699	Sequence 170699,
35	66.6	6.3	3957	16	US-10-200-562-193	Sequence 193, App
36	66.6	6.3	3957	16	US-10-237-551-193	Sequence 193, App
37	66.6	6.3	154746	10	US-09-827-688-8	Sequence 8, Appli
c	66.6	6.3	154746	10	US-09-827-688-8	Sequence 8, Appli
c	66.2	6.2	1614	9	US-09-976-740-45	Sequence 45, Appli
c	66.2	6.2	1614	13	US-10-023-529-45	Sequence 45, Appli
c	66.2	6.2	1614	13	US-10-023-523-45	Sequence 45, Appli
c	66.2	6.2	1614	17	US-10-616-187-45	Sequence 45, Appli
c	66.2	6.2	1614	17	US-10-671-242-45	Sequence 45, Appli
c	66.2	6.2	12425	9	US-09-976-740-50	Sequence 50, Appli
c	66.2	6.2	12425	13	US-10-023-529-50	Sequence 50, Appli

ALIGNMENTS

RESULT 1

US-10-672-399-7
; Sequence 7, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; PRIOR FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-7

Query Match	100.0%;	Score 1065;	DB 18;	Length 1065;
Best Local Similarity	100.0%;	Pred. No. 2.8e-271;		
Matches 1065;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	ATGAGACAGCAGGAGCGCCCAAGCCCACTCTCGAGCCCGCCGCTGCTCGGCGCTGGCC	60	
Db	1	ATGAGACAGCAGGAGCGCCCAAGCCCACTCTCGAGCCCGCCGCTGCTCGGCGCTGGCC	60	
Qy	61	CGGAGGGTGTGACCATCGCCTTCGCCCTGTCTCTCTGCGGCTCATGACCTGGGCGCTAC	120	
Db	61	CGGAGGGTGTGACCATCGCCTTCGCCCTGTCTCTCTGCGGCTCATGACCTGGGCGCTAC	120	
Qy	121	CCGCGCGGGGTGCGCGCTGCGCTCCGATCGCTACGCGCTTCCTGCGCTTCGCGCTCTACGGG	180	

Db 121 GCCCGCGGGTCCGCTCGATCGTACGGCTCTCGGCTTGGCCCTCTACGG 180
Qy 181 GCCTTCTTTAGCGCACTGTGGCGAGAGCTTTCGCTACTTGGAGCACCGGG 240
Db 181 GCCTTCTTTAGCGCACTGTGGCGAGAGCTTTCGCTACTTGGAGCACCGGG 240
Qy 241 GTGGCGGGCGGGCGGGCGGCTGTGATGACGACCGCGCGAGTGTGGCGCTGACC 300
Db 241 GTGGCGGGCGGGCGGGCGGCTGTGATGACGACCGCGCGAGTGTGGCGCTGACC 300
Qy 301 ATCTCCGCTACACAGAGACCCCGCTACCTGCGCCAGTGTGCTGGCGTCCGCGGCC 360
Db 301 ATCTCCGCTACACAGAGACCCCGCTACCTGCGCCAGTGTGCTGGCGTCCGCGGCC 360
Qy 361 CTGCTGTACCGCGCGGGCTGTGCTCTCATGTGGTGGATGGCAACCGCGCCGAG 420
Db 361 CTGCTGTACCGCGCGGGCTGTGCTCTCATGTGGTGGATGGCAACCGCGCCGAG 420
Qy 421 GACCTCTACATGGTGCAGATGTTCCGCGAGGTCTTTCGCTGACGAGACCCCGCCACGTAC 480
Db 421 GACCTCTACATGGTGCAGATGTTTCGCGAGGTCTTTCGCTGACGAGACCCCGCCACGTAC 480
Qy 481 GTGTGGGACGGCAATACACAGCCCTGGGAACCCCGCGCGGGCGGCTGGGGCGCC 540
Db 481 GTGTGGGACGGCAATACACAGCCCTGGGAACCCCGCGCGGGCGGCTGGGGCGCC 540
Qy 541 GGAGCTATCGGAGGTGGAGCGAGGATTCCTGGGGCGCTGGAGTGGAGCGCTGTG 600
Db 541 GGAGCTATCGGAGGTGGAGCGAGGATTCCTGGGGCGCTGGAGTGGAGCGCTGTG 600
Qy 601 AGGACTCCGAGTGTGCTGCGCAGCGCTGGGGCGGCAAGCGAGGTCTATGTAC 660
Db 601 AGGACTCCGAGTGTGCTGCGCAGCGCTGGGGCGGCAAGCGAGGTCTATGTAC 660
Qy 661 ACAGCTTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGTGACTCGGACACA 720
Db 661 ACAGCTTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGTGACTCGGACACA 720
Qy 721 AGGTTGAGACCCATGGCACTGTGAGCTCGTGGGGTACTGGAGAGACCCCGGGTA 780
Db 721 AGGTTGAGACCCATGGCACTGTGAGCTCGTGGGGTACTGGAGAGACCCCGGGTA 780
Qy 781 GGGGCTGTGTGGGAGCTGCGGATCTTAACCTCTTGACTCTGGGTGAGTTCCTA 840
Db 781 GGGGCTGTGTGGGAGCTGCGGATCTTAACCTCTTGACTCTGGGTGAGTTCCTA 840
Qy 841 AGCAGCTGCGATACTGGGTAGCTTCAATGTGGAGCGGGTGTGTGACTTCGACACA 900
Db 841 AGCAGCTGCGATACTGGGTAGCTTCAATGTGGAGCGGGTGTGTGACTTCGACACA 900
Qy 901 TGTGTATCTGATCAGCGGTCTCTAGGCTATATAGGATTAACCTCTTGGAGAGTTT 960
Db 901 TGTGTATCTGATCAGCGGTCTCTAGGCTATATAGGATTAACCTCTTGGAGAGTTT 960
Qy 961 CTTGAGGCTGTGTACCAACAGAGTTCCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1020
Db 961 CTTGAGGCTGTGTACCAACAGAGTTCCTGGGTACCCACTGTACTTTTGGGGATGACCGG 1020
Qy 1021 CACCTCACCAACCGCATGCTCAGCATGGGTATGTATACCAAGTAA 1065
Db 1021 CACCTCACCAACCGCATGCTCAGCATGGGTATGTATACCAAGTAA 1065

RESULT 2

US-10-672-399-1
; Sequence 1, Application US/10672399
; Publication No. US2005003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25

; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-1

Query Match 99.9%; Score 1064; DB 18; Length 1737;
Best Local Similarity 100.0%; Pred. No. 5.3e-271;
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGACGACGAGGACGGCCCAAGCCCACTCTCTGACGCCCGCCGCTGTCTGGGCTTGSCC 60
Db 1 ATGAGACGACGAGGACGGCCCAAGCCCACTCTCTGACGCCCGCCGCTGTCTGGGCTTGSCC 60
Qy 61 CGGAGGGTGTGACCATCGCCTTGCCTGTCTCATCTTGGGCCCTCATGACCTTGGSCCTTAC 120
Db 61 CGGAGGGTGTGACCATCGCCTTGCCTGTCTCATCTTGGGCCCTCATGACCTTGGSCCTTAC 120
Qy 121 GCCCGCGGGGTGCGGCTCGGATCGCTACGCGCTCTCTGGGCTTTCGGGCTCTTACGGG 180
Db 121 GCCCGCGGGGTGCGGCTCGGATCGCTACGCGCTCTCTGGGCTTTCGGGCTCTTACGGG 180
Qy 181 GCCTTCTTTTTCAGCGCACTGTGTGGCGAGAGCTCTTTCGGGTACTTGGAGACCGGGCGG 240
Db 181 GCCTTCTTTTTCAGCGCACTGTGTGGCGAGAGCTCTTTCGGGTACTTTCGGAGCACCGGGCGG 240
Qy 241 GTGGCGGGCGGGCGGGCGGCTGGATGCAGCCACCGCGCGAGTGTGGCGCTGAC 300
Db 241 GTGGCGGGCGGGCGGGCGGCTGGATGCAGCCACCGCGCGAGTGTGGCGCTGAC 300
Qy 301 ATCTCCGCTACACAGAGGACCCCGCGCTACCTGGCGCAAGTCTCTGGGCTCCGCGGCC 360
Db 301 ATCTCCGCTACACAGAGGACCCCGCGCTACCTGGCGCAAGTCTCTGGGCTCCGCGGCC 360
Qy 361 CTGCTGTATCCGCGCGGCTCGGCTCGGCTCTCATGTGTGGTGGATGGCAACCGGGCCGAG 420
Db 361 CTGCTGTATCCGCGCGGCTCGGCTCGGCTCTCATGTGTGGTGGATGGCAACCGGGCCGAG 420
Qy 421 GACCTCTACATGTGTGACATGTTCCGCGAGGTCTTTCGCTGACGAGACCCCGCCACGTAC 480
Db 421 GACCTCTACATGTGTGACATGTTCCGCGAGGTCTTTCGCTGACGAGACCCCGCCACGTAC 480
Qy 481 GTGTGGGACGGCAATACACAGCCCTGGGAACCCCGCGCGGGCGGCTGGGTGGGCGCC 540
Db 481 GTGTGGGACGGCAATACACAGCCCTGGGAACCCCGCGCGGGCGGCTGGGTGGGCGCC 540
Qy 541 GGAGCTATCGGAGGTGGAGCGAGGATTCCTGGGGCTGGAGTGGAGGCGCTGTG 600
Db 541 GGAGCTATCGGAGGTGGAGCGAGGATTCCTGGGGCTGGAGTGGAGGCGCTGTG 600
Qy 601 AGGACTCGAGTGTGCTGCGCAGCGCTGGGGCGGCAAGCGAGGTCTATGTAC 660
Db 601 AGGACTCGAGTGTGCTGCGCAGCGCTGGGGCGGCAAGCGAGGTCTATGTAC 660
Qy 661 ACAGCTTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGTGACTCGGACACA 720
Db 661 ACAGCTTTCAAGCGCTCGGAGATTCGGTGGACTACGTGCAAGTCTGTGACTCGGACACA 720
Qy 721 AGGTTGAGACCCATGGCACTGTGAGCTCGTGGGGTACTTGGAGAGACCCCGGGTA 780
Db 721 AGGTTGAGACCCATGGCACTGTGAGCTCGTGGGGTACTTGGAGAGACCCCGGGTA 780
Qy 781 GGGGCTGTGTGGGAGCTGCGGATCTTAACCTCTTGACTCTGGGTGAGTTCCTA 840
Db 781 GGGGCTGTGTGGGAGCTGCGGATCTTAACCTCTTGACTCTGGGTGAGTTCCTA 840
Qy 841 AGCAGCTGCGATACTGGGTAGCTTCAATGTGGAGCGGGTGTGTGACTTCGACACA 900
Db 841 AGCAGCTGCGATACTGGGTAGCTTCAATGTGGAGCGGGTGTGTGACTTCGACACA 900

QY 901 TGTGTATCTGTCATCAGCGGTCCTCTAGGCTTATATAGGAATTAACCTCTTTCAGCAGTTT 960
Db 901 TGTGTATCTGTCATCAGCGGTCCTCTAGGCTTATATAGGAATTAACCTCTTTCAGCAGTTT 960
QY 961 CTGAGGCTGTGTACAAACAGAGTCTCTGGTACCCACTGTACTTTTGGGATGACCGG 1020
Db 961 CTGAGGCTGTGTACAAACAGAGTCTCTGGTACCCACTGTACTTTTGGGATGACCGG 1020
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGTTATGTATACCAAGTA 1064
Db 1021 CACCTCACCAACCGCATGCTCAGCATGGTTATGTATACCAAGTA 1064

RESULT 3

US-10-042-523-1
; Sequence 1, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/042,523
; FILING DATE: 19-Oct-2001
; CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A

FILING DATE: 22-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Brook, David E.

REGISTRATION NUMBER: 22,592

REFERENCE/DOCKET NUMBER: LKS95-07

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2116 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 36..1769

SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-042-523-1

Query Match 99.9%; Score 1064; DB 13; Length 2116;
Best Local Similarity 100.0%; Pred. No. 5.3e-271; Indels 0; Gaps 0;
Matches 1064; Conservative 0; Mismatches 0;

QY 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCCCGCGCTGTCTCGGCTGGCC 60
Db 36 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCCCGCGCTGTCTCGGCTGGCC 95
QY 61 CGGAGGCTGTGACCATCGCTTCGCGCTTCGCTATCTCTGGGCTCATGACCTGGGCTTAC 120
Db 96 CGGAGGCTGTGACCATCGCTTCGCGCTTCGCTATCTCTGGGCTCATGACCTGGGCTTAC 155

QY 121 GCCGCGGGGTGCGGCTGGCTCCGATCGCTACGGCTCTCTGGCTTCGGCTCTACGGG 180
Db 156 GCCGCGGGGTGCGGCTGGCTCCGATCGCTACGGCTCTCTGGCTTCGGCTCTACGGG 215
QY 181 GCCTCTCTTTTTCAGGCGACCTGGTGGCGGAGAGCTTCTTCGGGTACTGAGAGCAGCGGG 240
Db 216 GCCTCTCTTTTTCAGGCGACCTGGTGGCGGAGAGCTTCTTCGGGTACTGAGAGCAGCGGG 275
QY 241 GTGGCGGGCGGGCGGGGGCGGCTGATGACGACACCGCGCGCAGTGTGGCGCTGACC 300
Db 276 GTGGCGGGCGGGCGGGGGCGGCTGATGACGACACCGCGCGCAGTGTGGCGCTGACC 335
QY 301 ATCTCCGCTTACCAGGAGGACCCCGCTGACCTGCGCCAGTGCCTGGCGTCCGCCCCGGCC 360
Db 336 ATCTCCGCTTACCAGGAGGACCCCGCTGACCTGCGCCAGTGCCTGGCGTCCGCCCCGGCC 395
QY 361 CTGCTGTACCCGCGGGCGGGCTGGCGCTCTCATGTGGTGGATGGCAACCGCGCCGAG 420
Db 396 CTGCTGTACCCGCGGGCGGGCTGGCGCTCTCATGTGGTGGATGGCAACCGCGCCGAG 455
QY 421 GACCTCTACATGCTGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGTAC 480
Db 456 GACCTCTACATGCTGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGTAC 515
QY 481 GTGTGGGACGGCAACTACACAGCCCTTGGGAACCCCGCGGGCGGGCGGCTGGCGGCC 540
Db 516 GTGTGGGACGGCAACTACACAGCCCTTGGGAACCCCGCGGGCGGGCGGCTGGCGGCC 575
QY 541 GGAGCCTATCGGAGGTGGAGGGGAGGATCTTGGGGCGGCTGGCAGTGGAGCGCTGGTG 600
Db 576 GGAGCCTATCGGAGGTGGAGGGGAGGATCTTGGGGCGGCTGGCAGTGGAGCGCTGGTG 635
QY 601 AGGACTCCAGGTGGTGGTGGCGCAGCGCTGGGGCGGCAAGCGGAGGTCTATGTAC 660
Db 636 AGGACTCCAGGTGGTGGTGGCGCAGCGCTGGGGCGGCAAGCGGAGGTCTATGTAC 695
QY 661 ACAGCCTTCAAGGGCTCGGAGATTTCGGTGGATCTACGTGCAGGTCTGTGACTCGGACACA 720
Db 696 ACAGCCTTCAAGGGCTCGGAGATTTCGGTGGATCTACGTGCAGGTCTGTGACTCGGACACA 755
QY 721 AGTTGGACCCCATGGGCACTGCTGGAGCTCTGGGGGTACTGGAGCAGAGACCCCGGGTA 780
Db 756 AGTTGGACCCCATGGGCACTGCTGGAGCTCTGGGGGTACTGGAGCAGAGACCCCGGGTA 815
QY 781 GGGGCTGTGGTGGGAGCTGGGATCCTTAACCTCTGGACTCTCTGGGTCTGAGCTTCCTTA 840
Db 816 GGGGCTGTGGTGGGAGCTGGGATCCTTAACCTCTGGACTCTCTGGGTCTGAGCTTCCTTA 875
QY 841 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCAC 900
Db 876 AGCAGCCTGCGATACTGGGTAGCCTTCAATGTGGAGCGGGCTTGTGAGAGTACTTCCAC 935
QY 901 TGTGTATCTCGATCAGCGGCTCTTAGGCTTATAGGAATAACCTCTTTCAGCAGTTT 960
Db 936 TGTGTATCTCGATCAGCGGCTCTTAGGCTTATAGGAATAACCTCTTTCAGCAGTTT 995
QY 961 CTGAGGCTGTGACAAACAGAGTTCCTGGGTACCCACTGACTTTTGGGATGACCGG 1020
Db 996 CTGAGGCTGTGACAAACAGAGTTCCTGGGTACCCACTGACTTTTGGGATGACCGG 1055
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGTTATGTACCAAGTA 1064
Db 1056 CACCTCACCAACCGCATGCTCAGCATGGTTATGTACCAAGTA 1099

RESULT 4

US-10-672-399-5
; Sequence 5, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics

```
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-5

Query Match      87.1%; Score 928; DB 18; Length 1083;
Best Local Similarity 100.0%; Pred. No. 4.1e-235;
Matches 928; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ATGAGACAGCAGGAGCGCCCAAGCCCACTCTCGCAGCGCCGCGCTGCTCGGCGCTGACC 60
DB      1 ATGAGACAGCAGGAGCGCCCAAGCCCACTCTCGCAGCGCCGCGCTGCTCGGCGCTGACC 60

QY      61 CGGAGGCTGCTGACCATCGCTTCCGATCGCTACGCGCTCTCGGCGCTGACCGGCTAC 120
DB      61 CGGAGGCTGCTGACCATCGCTTCCGATCGCTACGCGCTCTCGGCGCTGACCGGCTAC 120

QY      121 GCCCGCGGGGTGCGCGCTCGATCGCTACGCGCTCTCGGCGCTGACCGGCTACCGG 180
DB      121 GCCCGCGGGGTGCGCGCTCGATCGCTACGCGCTCTCGGCGCTGACCGGCTACCGG 180

QY      181 GCCTTCCTTTAGCGCACTGCTGAGCGCGCGCTGATGCGCGCGCGCTGAGCGCGCTGAC 240
DB      181 GCCTTCCTTTAGCGCACTGCTGAGCGCGCGCTGATGCGCGCGCGCTGAGCGCGCTGAC 240

QY      241 GTGCGCGCGCGCGCGCGCGCGCGCTGATGCGCGCGCGCGCTGAGCGCGCGCTGAC 300
DB      241 GTGCGCGCGCGCGCGCGCGCGCGCTGATGCGCGCGCGCGCTGAGCGCGCGCTGAC 300

QY      301 ATCTCCGCTTACAGGAGCGCGCGCGCTGATGCGCGCGCGCGCTGAGCGCGCGCTGAC 360
DB      301 ATCTCCGCTTACAGGAGCGCGCGCGCTGATGCGCGCGCGCGCTGAGCGCGCGCTGAC 360

QY      361 CTGCTGTACATGCTGCAATGCTTCCGAGGCTTCTCGCTGACGAGGACCCCGCGCGCGCTGAC 420
DB      421 GACCTCTACATGCTGCAATGCTTCCGAGGCTTCTCGCTGACGAGGACCCCGCGCGCGCTGAC 420

QY      481 GTGTGGAGCGCACTACCAAGCCCTTGGGAAACCCCGCGCGCGCGCGCGCTGAGCGCGCTGAC 540
DB      481 GTGTGGAGCGCACTACCAAGCCCTTGGGAAACCCCGCGCGCGCGCGCGCTGAGCGCGCTGAC 540

QY      541 GGAGCTATCGGAGGCTGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCT 600
DB      541 GGAGCTATCGGAGGCTGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCT 600

; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-3

; Sequence 3, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-3

Query Match      87.1%; Score 927.4; DB 18; Length 1071;
Best Local Similarity 99.9%; Pred. No. 5.9e-235;
Matches 928; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ATGAGACAGCAGGAGCGCGCCCAAGCCCACTCTCGCAGCGCCGCGCTGCTCGGCGCTGACC 60
DB      1 ATGAGACAGCAGGAGCGCGCCCAAGCCCACTCTCGCAGCGCCGCGCTGCTCGGCGCTGACC 60

QY      61 CGGAGGCTGCTGACCATCGCTTCCGCTGCTCATCTCGGCGCTCATGACCTGCGGCTTAC 120
DB      61 CGGAGGCTGCTGACCATCGCTTCCGCTGCTCATCTCGGCGCTCATGACCTGCGGCTTAC 120

QY      121 GCCCGCGGGGTGCGCGCTCGATCGCTTACGCGCTCTCGGCGCTTCTCGGCGCTTCTACCGG 180
DB      121 GCCCGCGGGGTGCGCGCTCGATCGCTTACGCGCTCTCGGCGCTTCTCGGCGCTTCTACCGG 180

QY      181 GCCTTCCTTTAGCGCACTGCTGAGCGCGCGCTGATGCGCGCGCGCTGAGCGCGCGCTGAC 240
DB      181 GCCTTCCTTTAGCGCACTGCTGAGCGCGCGCTGATGCGCGCGCGCTGAGCGCGCGCTGAC 240

QY      241 GTGCGCGCGCGCGCGCGCGCGCGCTGATGCGCGCGCGCGCTGAGCGCGCGCTGACC 300
DB      241 GTGCGCGCGCGCGCGCGCGCGCGCTGATGCGCGCGCGCGCTGAGCGCGCGCTGACC 300

QY      301 ATCTCCGCTTACAGGAGGAGCCCGCGCTGATGCGCGCGCGCTGAGCGCGCGCTGACC 360
DB      301 ATCTCCGCTTACAGGAGGAGCCCGCGCTGATGCGCGCGCGCTGAGCGCGCGCTGACC 360

QY      361 CTGCTGTATACCGCGCGCGCGCTGCGCTGCTCATATGCTGCTGATGCGCAACCGGCGCGGAG 420
DB      361 CTGCTGTATACCGCGCGCGCGCTGCGCTGCTCATATGCTGCTGATGCGCAACCGGCGCGGAG 420

QY      421 GACCTCTACATGCTGCAATGCTTCCGAGGCTTCTCGCTGACGAGGACCCCGCGCGCGCTGAC 480
DB      421 GACCTCTACATGCTGCAATGCTTCCGAGGCTTCTCGCTGACGAGGACCCCGCGCGCGCTGAC 480

QY      481 GTGTGGAGCGCACTACCAAGCCCTTGGGAAACCCCGCGCGCGCGCGCGCTGAGCGCGCTGAC 540
DB      481 GTGTGGAGCGCACTACCAAGCCCTTGGGAAACCCCGCGCGCGCGCGCGCTGAGCGCGCTGAC 540

QY      541 GGAGCTATCGGAGGCTGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCT 600
DB      541 GGAGCTATCGGAGGCTGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCT 600
```



```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-27210

Query Match 39.5%; Score 421; DB 10; Length 490;
Best Local Similarity 90.9%; Pred. No. 2.6e-101;
Matches 442; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 544 GCCTATCGGAGGTGAGCGGAGGATCTTGGCGGCTGGCAGTGAGGCGCTGGTGAGG 603
DB 2 CGAGTAGGGANNNNNNAATTGAGACTGAGGCACNGTCGGGAANTCGCTGGTGAGG 61
QY 604 ACTCGCAGGTGCTGTGGTGGCGCAGCGCTGGGCGCGCAAGCGAGGTGATGTACACA 663
DB 62 ACTCGCAGGTGCTGTGGTGGCGCAGCGCTGGGCGCGCAAGCGGTGATGTACACA 121
QY 664 GCCTTCAAGCGCTCGGAGATTGCGTGACGTACGTGCAAGGTCTGTGACTCGGACACAAGG 723
DB 122 GCCTTCAAGCGCTCGGAGATTGCGTGACGTACGTGCAAGGTCTGTGACTCGGACACAAGG 181
QY 724 TTGGACCCCATGGCACTCTGGAGCTCTGGGCTGCTGGGCTGCTGGGCTGCTGGGCTGCTGGG 783
DB 182 TTGGACCCCATGGCACTCTGGAGCTCTGGGCTGCTGGGCTGCTGGGCTGCTGGGCTGCTGGG 241
QY 784 GCTGTGTGGGAGCGTGGGATCTTAAACCTCTGGGCTCTGGGCTGCTGGGCTCTGGGCTCTGGG 843
DB 242 GCTGTGTGGGAGTGGGATCTTAAACCTCTGGGCTCTGGGCTCTGGGCTCTGGGCTCTGGGCT 301
QY 844 AGCTTGCATCTGGGTAGCTTCAATGTGAGCGGGCTTGTGAGAGTCTTGTGAGAGTCTTGTGAG 903
DB 302 AGCTTGCATCTGGGTAGCTTCAATGTGAGCGGGCTTGTGAGAGTCTTGTGAGAGTCTTGTGAG 361
QY 904 GTATCTTCATCAGCGGCTCTAGGCTTATAGGATACCTCTTGCAGGAGTTTCTT 963
DB 362 GTATCTTCATCAGCGGCTCTAGGCTTATAGGATACCTCTTGCAGGAGTTTCTT 421
QY 964 GAGGCTTGTACACAGAGTTCTGGGTACCCACTGCTACTTTTGGGGATGACCGGCAC 1023
DB 422 GAGGCTTGTACACAGAGTTTCTGGGTACCCACTGCTACTTTTGGGGATGACCGGCCT 481
QY 1024 CTCACC 1029
DB 482 CACAAC 487

RESULT 10
US-10-363-345A-13973/c
; Sequence 13973, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 13973
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-Island No: 13973
US-10-363-345A-13973

Query Match 37.0%; Score 393.8; DB 18; Length 662;
Best Local Similarity 74.7%; Pred. No. 4e-94;
Matches 494; Conservative 0; Mismatches 167; Indels 0; Gaps 0;
```

```
QY 335 CAGCCCGCGCTGCTCCGGCTTGGCCCGAGGGTGTGACCATCGCCTTGGCCCTGCTCA 94
DB 661 CAACCTACCGCTACTCCGACCTAACCCGAAATACTAACCATCGCCTTGGCCCTTACTCA 602
QY 95 TCCTGGGCTCATGACCTGGGCTTACGCGCGCGGGTGCCTGCTCCGATCGCTAGG 154
DB 601 TCCTAAACCTCATAACTTAAACCTTACGCGCGCGGAATAACCGCTAACCTCCGATCGCTAGG 542
QY 155 GCCTCTGGCTTGGGCTTACGCGGCTTCTTTCAGCGCACCTGGTGGCGAGAGCC 214
DB 541 ACCTCTAACCTTGGACCTTACGAAACCTTCTTTCACCGCACCTTAATAACGCAAAACC 482
QY 215 TCTTCGCTACTTGGAGCACCGCGGGTGGCGCGCGCGCGCGCGCGCTGGATGCGAG 274
DB 481 TCTTCGCTACTTAAACACACCGAGAAATAACGAGACGACGCGAAATACGCTAAATACAA 422
QY 275 CCACCGCGCAGTGTGGCGCTGACCATCTCGCTTACAGGAGGACCCCGGCTACCTGC 334
DB 421 CCACCGCGCAATATATACGCTAACCATCTCGCTTACCAAAACCCCGGCTACCTAC 362
QY 335 GCCAGTCTGCGCTCGCGCGCGCGCTGTGTACCGCGCGCGCGCTGCGCGCTCTCA 394
DB 361 GCCAATACCTAACGCTCGCGCGCGCTTATATACCGCGCGCGCTTACCGCTCTCA 302
QY 395 TGGTGTGGATGGCAACCGCGCGCGCGCTTCTACATGCTCGACATGTTCCGCGAGGTCT 454
DB 301 TAATAATAATAACCAACCGCGCGCGCTTCTACATAATCGACATATTCGCGGAAATCT 242
QY 455 TCGCTGACGAGGACCCCGCGCTACGTGTGGGAGCGGCAATACACAGCCCTGGGAAC 514
DB 241 TCGCTAACGAAACCCCGCGCTACGTGTATAAACGCAACATACCAACCAACCTTAAAC 182
QY 515 CCGCGCGCGCGCGCGCTGGGCGCGCGCTTACCGAGGCTTACCGAGGCTGGAGGAGTCCCTG 574
DB 181 CCGCGACGACGAGCGGATAAACCGCGGAAACCTTATCGAAATAATAAAACGAAATAATCTTA 122
QY 575 GCGCGCTGGAGTGGAGCGCTGGTGGAGGCTTGGAGGCTTGGCGTGGCGGAGCGCT 634
DB 121 AACGACTAACATAAAACCGCTTAAATAACCTCGCAATACGTATACGTAACGCAACGCT 62
QY 635 GGGCGGCAAGCGGAGGTGATGTACAGGCTTCAAGCGCTCGGAGGCTCGGATGCGTGGACT 694
DB 61 AAAACGACAAACGCGAATCATATACACACCTTCAAAACGCTCGAAATTCGATAAACT 2
QY 695 A 695
DB 1 A 1
```

```
RESULT 11
US-10-363-345A-13974
; Sequence 13974, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT APPLICATION NUMBER: US/10/363,345A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 13974
; LENGTH: 662
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: Cpg-Island No: 13974
US-10-363-345A-13974
```

```
Query Match      37.0%; Score 393.8; DB 18; Length 662;
Best Local Similarity 74.7%; Pred. No. 4e-94;
Matches 494; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

Qy 35 CAGCCCGCGCTGCTCCGCGCTTGGCGGAGGTGCTGACCATGCGCTTGGCGCTGCTCA 94
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
2 CAACCTACCGCTACTCCGACCTTAACCCGAAATACTAACCATCGCTTCCGCTTACTCA 61

Qy 95 TCCTGGGCTCATGACCTGGGCTACGCGCGGGGTGCGGCTGCGCTCCGATCGCTACG 154
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
62 TCCTAAACCTATAACCTAAACCTACGCGCGGAAATACCGCTAACCTCCGATCGCTACG 121

Qy 155 GCCTCTCGGCTTCCGCTCTACGCGGCTTCTCTTTCAGCGCACCTGGTGGCGAGAGCC 214
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
122 ACCTCTAACCTTCACTTACGAACTTCTCTTTCAGCGCACCTTAACGAAACCAAC 181

Qy 215 TCTTCGCTACTGGAGCACCGCGGGTGGCGGCGGCGGGGGCGCTGGATGAG 274
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
182 TCTTCGCTACTAAACACCGACGAATAACGACGACGCGAAACCGCTAAATACAA 241

Qy 275 CCACGCGCGAGTGGGCTGACCATCTCGCTTACCGCGCGCGGCTGCGCTCTCTCA 334
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
242 CCACGCGCGCAATATAACGCTAACCATCTCGCTTACCAAAACCCCGCTACCTAC 301

Qy 335 GCCAGTCTGCGCTCCGCGCGCTGCTGTATCCCGCGCGCGGCTGCGCTCTCTCA 394
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
302 GCCAATACCTAACGCTCCGCGCGCTTACTATACCGCGCGCGGCTACGCGTCTCA 361

Qy 395 TGGTGGTGGATGGCAACCGCGCGAGGACCTCTATCATGGTGCACTGTTCCGCGAGTCT 454
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
362 TAATAATAATAACCAACCGCGCGGAAACCTCTTACATAATGACATATTCGCGGAAATCT 421

Qy 455 TCGCTGACGAGGACCCCGCACGTAAGTGGGAGCGCACTACACACGCGCTGGGAAC 514
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
422 TCGCTAACGAAACCCCGCACGTAAGTGGGAGCGCACTACACACGCGCTGGGAAC 481

Qy 515 CCGCGCGCGCGCGCGCTGGCGCGCGGAGCTATCGGAGGTGGAGCGGAGATCTG 574
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
482 CCGGACGACGACGAGTAACCGCGGAAACCTATCGAAATAATAAACGAAATCTTA 541

Qy 575 GCGGCTGGCAGTGGAGCGCTGGTGGAGACTGCGAGTGGCTGGCTGGCGGAGCGCT 634
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
542 AACGACTAACCAATAAAACCGCTTAATAAAACCTCGCAATAGCTATACGTAACGCAACGCT 601

Qy 635 GGGCGGCAACGCGGAGTCTATGATACAGGCTTCAAGCGCTCGGAGATTCGTTGACT 694
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
602 AAAACGACAAACGCGAAATCATATACAAACCTTCAAAACGCTCGAAATTCGATAAAT 661

Qy 695 A 695
Db 662 A 662
```

RESULT 12
US-10-295-027-369
; Sequence 369, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13

```
; PRIOR APPLICATION NUMBER: US 09/663,733  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/335,394  
; PRIOR FILING DATE: 2001-11-15  
; PRIOR APPLICATION NUMBER: US 60/332,464  
; PRIOR FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: US 60/334,393  
; PRIOR FILING DATE: 2001-11-29  
; PRIOR APPLICATION NUMBER: US 60/340,376  
; PRIOR FILING DATE: 2001-12-14  
; PRIOR APPLICATION NUMBER: US 60/347,211  
; PRIOR FILING DATE: 2002-01-08  
; PRIOR APPLICATION NUMBER: US 60/347,349  
; PRIOR FILING DATE: 2002-01-10  
; PRIOR APPLICATION NUMBER: US 60/355,250  
; PRIOR FILING DATE: 2002-02-08  
; PRIOR APPLICATION NUMBER: US 60/356,714  
; PRIOR FILING DATE: 2002-02-13  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 1386  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 369  
; LENGTH: 1662  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-295-027-369
```

```
Query Match      33.1%; Score 353; DB 17; Length 1662;  
Best Local Similarity 62.9%; Pred. No. 2.4e-83;  
Matches 633; Conservative 0; Mismatches 335; Indels 39; Gaps 4;  
Qy 61 CGGAGGTGTGACCATCGCTTCCGCTTCTCTCTTGGGCTTACGCTTGGGCTTAC 120  
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
28 CGTGTGGTGGCACACGACCTGTTTGGCTTGGGCTTGGGCTTGGGCTTGGGCTTAT 87  
Qy 121 GCGCGCGGGTGGCTTGGCTTCCGCTTACGCTTACGCTTCTTGGCTTCTTACGGG 180  
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
88 GTGACGGGCTTACGCTTCTTACGCTTCTTACGCTTCTTACGCTTCTTACGCTT 147  
Qy 181 GCCTTCTTTCAGCGCTTGGTGGCGAGAGCTTCTTCTTGGCTTCTTGGAGAGCGGCGG 240  
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
148 GCCATCTGGGCTTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 207  
Qy 241 GTGCGGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 297  
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
208 ATGCGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 267  
Qy 298 ACCATCTCGCTTACGAGGAGCGGCGGCTTACGCTTGGGCTTGGGCTTGGGCTTGGG 357  
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
268 TGCATTTGGGCTTACGAGGAGCGGCTTACGCTTGGGCTTGGGCTTGGGCTTGGGCT 327  
Qy 358 GCGCTGCTTACCGCGGCGGCGGCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 417  
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
328 CGCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 381  
Qy 418 GAGGACCTTACATGCTGAGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 477  
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
382 GAGGACGCTTACATGCTGAGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 441  
Qy 478 TACGTGTGGAGCGGCAACTTACACGAGCTTGGGAAACCGCGGCGGCGGCGGCGGCGGCGG 537  
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
442 TCTTGTGTGGCGAGCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 486  
Qy 538 GCGGAGCTTATCGGAGGTGGAGCGGAGGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 597  
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
487 ACAGGAGCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 531  
Qy 598 GTGAGGAGCTCGAGGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 657  
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||  
532 GTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 591
```

Qy	658	TACACAGCCTTCAAGGGCGCTCGGAGATTCCGTGTGACTACGTGCAAGTCTGTGACTCGGAC	717
Db	592	TACACGGCCTTCAAGAGGCCCTCGGCGATTCCGTGTGACTACATCCAGGTGTGCGACTCTGCAC	651
Qy	718	ACRAAGTTTGGACCCCATGGCACTCTCGAGCTCTGTGCGGGTACTGGACGAGGACCCCCGG	777
Db	652	ACTGTCTTGGATCCAGCCTGCAACCATCGAGATGCTTCCAGTCTCTGGAGGAGATCCCCAA	711
Qy	778	GTACGGGCTGTGTGGTGGGGAGCTGCGGATCCTTAACCTCTCTGGACTCCTCGGTGAGCTTC	837
Db	712	GTAGGGGGAGTGGGGGAGATGTCAGATCTCTCAACAAGTACGACTCATGGAATTTCTCTTC	771
Qy	838	CTAAGACGCGTGCATACTCGGTGAGCCTTCAATGTGGAGCGGGCTTGTGACAGCTACTTC	897
Db	772	CTGACGACGCTGCGGTACTGGAATGGCCCTTCAACGTGGAGCGGGCTCCGACGTCTACTTTT	831
Qy	898	CACGTGTATCTTCGCATCAGCGGTCTCTAGCGCTATATAGGAATAACCTCTTTCACGACG	957
Db	832	GGCTGTGTGCAGTGTATTAGTGGGCGCTTTGGGCATGTACCGCAACAGCTCCTCCACGACG	891
Qy	958	TTTCTTGGAGGCCTGGTACAACCAGAAGTTCCTCGGGTATCCCACTGTACTTTTGGGGATGCAC	1017
Db	892	TTCTCTGGAGGACTGTGTACCATCAGAAGTTCCTAGGCAGCAAGTGCAGCTTCGGGGATGCAC	951
Qy	1018	CGGCACCTCACCAACCGGATGCTCAGCATGGGTTATGCTACCAAGTA	1064
Db	952	CGGCACCTCACCAACCGAGTCTCAGCGCTTTGGCTTACCGAACTAAGTA	998

RESULT 13

```

US-10-188-832-136
; Sequence 136, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: And Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188.832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 136
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-188-832-136

```

QY	181	GCCTTCCTTTACGCGCACTTGGTGGCGCAGAGCCTCTTCGCGTACTCTGGAGCAACCGCGCG	240
DB	148	GCCATCCTGGCGCCTGCACCTGCTCAATTCAGAGCCCTTTTGCCTTCTCTGGAGCAACCGCGCG	207
QY	241	GTGCGCGCGCGCGCGCGCGCGCGCTGCGTGCAGTGCAGCCACCGGC	297
DB	208	ATGCGACGTGCGCGCGCAGGCGCCCTGAAGCTGCCTCCCGCGCGGCGCTCGTGGCGACTG	267
QY	298	ACCATCTCCGCTACCAAGGAGGACCCCGCGTACTGCGCCAGTGCCTCTGGCGGTCCGCGCGCG	357
DB	268	TGCATTGCGCGGTACCAAGGAGGACCTGTACTATTGCGCAAGTGCCTGCGCTCGGCCCGAG	327
QY	358	GCCCTGTGTACCGCGCGCGCGCGCTGCGCTCTCATGTGTGTGATGGAACACCGCGCC	417
DB	328	CGCATCTCCTTCCCTGAC-----CTCAAGTGGTCAATGCTGGTGGATGCGCAACCGCGCAG	381
QY	418	GAGGACCTTACATGGTGCACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCGCAG	477
DB	382	GAGGACCCCTACATGCTTGGACATCTTCCACGAGGTGCTGGCGCGCACCGGAGCGCGCG	441
QY	478	TACGTGTGGAGCGCAACTACCAACCGCCCTGGGAACCCGCGCGCGCGCGCGGTGGCG	537
DB	442	TTCTTTGTGTGGCGCACAATTCCA-----TGAGGCGAGCGAGGGTGTAG	486
QY	538	GCCGCGAGCCTACTCGGAGGTTGGAGCGCGAGGATCTCTGGCGGCTGGCAGTGGAGCGCGCTG	597
DB	487	ACGAGGCCAGCCTGCGAGGAGGCGATGACCGT-----GTGCGGATGTG	531
QY	598	GTGAGGACTCGCAGGTGCGTGTGGTGGCGGAGCGGTGGGGCGGCAAGCGCGAGGTGATG	657
DB	532	GTGCGGCGCAGCACCTTCTCGTGCATCATGCAGAAAGTGGGAGGAGCAAGCGCGAGGTGATG	591
QY	658	TACACGCTTCAAGCGCTCGGAGATTCCGTGGACTACGTGCAGGTCTGTGACTCGGAC	717
DB	592	TACACGCGCTTCAAGGCGCTCGGCGATTCTGGTGGACTACATCCAGGTGTGCGACTCTGAC	651
QY	718	ACAAGGTTGGAACCCATCGCACTGCTGGAGCTGCTGGGGTACTTGGACGAGGACCCCGG	777
DB	652	ACTGTGTGATCCAGCCTGCACCATCGAGATGCTTCGAGTCTCTGGAGGAGGATCCCGAA	711
QY	778	GTAGGCGCTGTGTGGTGGGACGTGCGGATCCTTAACCTCTGGAAGTCTCTGGGTGAGTTTC	837
DB	712	GTAGGGCGAGTGGGGGAGATGTCCAGATCCTCAACAAGTACGACTCATGGATTTCTCTTC	771
QY	838	CTAAGCAGCCTCGGATACTGGGTAGCCTTCAATGTGGAGCGGCGTGTCTAGAGTACTTC	897
DB	772	CTGAGCAGCGTGGGTACTGGATGGCCCTCAACGTGGAGCGGCGCTGCCAGTCTCTACTTT	831
QY	898	CACGTGTATCTCTGCATCAGCGGTCTCTAGGCGCTATATAGGAATAACCTCTTGGACGAG	957
DB	832	GGCTGTGTGAGTGTATTAGTGGGCCCTTGGGATGTACCGCAACAGCCTCTCCAGCAG	891
QY	958	TTTCTTGAGCGCTGGTACAAACCGAAGTCTCTGGGTACCCACTGTACTTTTGGGGATGAC	1017
DB	892	TTCTGTGAGGACTGGTACCATCATAGAAGTTCTTAGGCGACAAGTGCAGTCTCGGGATGAC	951
QY	1018	CGGCACTCACCACCGCATGCTCAGCATGGGTTATGCTACCAAGTA	1064
DB	952	CGGCACCTCACCACCGAGTCTGTGAGCCTTGGGTACCGAACTAAGTA	998

RESULT 14

/ US-10-309-560-9
 / Sequence 9, Application US/10309560
 / Publication No. US20030235893A1
 / GENERAL INFORMATION:
 / APPLICANT: WEIGEL, PAUL H
 / APPLICANT: KUMARI, KSHAMA
 / TITLE OF INVENTION: HYALURONAN SYNTHASES AND METHODS OF MAKING AND USING SAME
 / FILE REFERENCE: 35341.082
 / CURRENT APPLICATION NUMBER: US/10/309,560
 / CURRENT FILING DATE: 2002-12-03
 / PRIOR APPLICATION NUMBER: 60/336,105

```

; PRIOR FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 9
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Xenopus laevis
US-10-309-560-9

```

Query Match	30.4%	Score 323.8	DB 17	Length 1767
Best Local Similarity	58.5%	Pred. No. 1.2e-75		
Matches	587	Conservative 0	Mismatches 407	Indels 9
				Gaps 1
Qy	62	GGAGGFGCTGACCATCGCCTTCGCCCTCGCTCATCTCTGGGCTCATGACCTGGGCGCTACG	121	
Db	80	GGAGGATAATTATATTTCTTTGGTGGTGGCTATTAGCTACCAATTACAGACGCTATG	139	
Qy	122	CCGCGGGGTGCGCTCGGCTCCGATCGCTACGGGCTCTGGCCCTTGGCCCTCTACGGGG	181	
Db	140	TGGCAGAGTCCAGGTCTCAAAATGAAGCAATCTCTCTCTCCCTGGGCTTTATGGTC	199	
Qy	182	CTTCTCTTTACGCGACCTGGTGGGCGAGAGCCTCTTCGCGTACCTGGAGCACCGGGGG	241	
Db	200	TTGCAATGCTTCTCCACGTGATGATGAGAGCCTCTTTGGCTTCTCTGGAGATAGCGAGG	259	
Qy	242	TGGCGGGCGGCGCGGGGCGCTGGATGACGACCAACCGCGCGAGTGTGGCGCTGACCA	301	
Db	260	TAAATAAGATGA-----GCTCTTCGAGCTTTAAGAAGACAGTGGCTCTGACCA	310	
Qy	302	TCTCGGCTTACAGAGAGACCCCGGTACCTGCGCCAGTGGCTGGCGTCCGCGCGCGCC	361	
Db	311	TTGCTGGGTATCAGAGAACCTTGATACCTGATAAAGTGTCTGGAATCTTGCAAGTATG	370	
Qy	362	TGCTGTATCCCGCGCGCGGCTGCGGTCTCATTGGTGGTGGATGGCAACCGCGCGGAGG	421	
Db	371	TGAAATACCCCAAGATAAACTCAAGATCATTTTGGTCATCGATGGGAAACAAGAGGATG	430	
Qy	422	ACCTCTACATGCTGCATGTTCCGCGAGGCTTCCTGCTGACGAGGACCCCGCCACGTACG	481	
Db	431	ATGCTTACATGATGGAGATGTTTCAAGACGTTTCCACGGTGAAGATGTAGGCACCTACG	490	
Qy	482	TGTGGGACGGCAACTACACACAGCCCTGGGAACCGCGCGCGGGCGGCTGGGCGCGCG	541	
Db	491	TATGGAAGGAAATTAACACACTGTATAAAGCTTGAGGACCAATAAGGATCCTGTCT	550	
Qy	542	GAGCCTATCGGAGGTGGAGCGGAGGATCTCTGGGCGGCTGGCAGTGGAGGCGCTGTGTGA	601	
Db	551	CTGAGGTTTCTTAAGCCCTTGAATGAAGATGAAGGTATCAATATGTTGGGAAGAACTTGTTA	610	
Qy	602	GGACTCGCAGTGGCTGGTGGCGCAGCGCTCGGGCGGCGCAAGCGGAGGTCAATGTACA	661	
Db	611	GAACAAGAGATGTGTGCATCATGCAACAGTGGGCGGGAAGAAGAGAGGTCAATGTACA	670	
Qy	662	CAGCCTTTCAAGGCGCTCGGAGATTCGGTGGACTACGTGCAGGTCTGTGACTCGGACACAA	721	
Db	671	CAGCATTTCCAGGCCATTTGGGACTTCTGTGGACTATGTACAGTCTGTGATCTCGACACCA	730	
Qy	722	GTTTGGACCCCATGCGCATCTGTGAGTCTGTGCGGGTACTGGACGAGAACCCCGGGTAG	781	
Db	731	AACTGGATGAATCGGCAACATGGAATTGGTGAAGGTTCTGGAATCCAATGACATGTATCG	790	
Qy	782	GGGCTGTTGGTGGGACGTGGGATCTTAACCTCTGGACTCTGGGTGAGCTTCCTAA	841	
Db	791	GGCAGTGGGAGGAGAGCTTCGCATCTGAACCCCTATGATTCCTCATTAGTTCATGA	850	
Qy	842	GCAGCTTGCATATCTGGGTAGCCTTCAATGTGAGCGGGCTTGTTCAGAGCTACTTCCACT	901	
Db	851	GCAGCTGCGTTACTGATGGCGTTTAACTGGAGAGGGCTGCCAGTCTTACTTCCACT	910	
Qy	902	GTGTATCTCTGCATCAGCGGTCCTCTAGGCCCTATATAGGAATAACCTCTTGACAGATTTTC	961	
Db	911	CGGTGCTCTGTATAAGTGGACCTCTGGGAATGTACCGGAACAACATCTCCAGGTGTTTT	970	

QY	962	TTGAGGCTTGGTACACACAGAGTTCTCTGGGTACCCACTGTACTTTTGGGGATGACCGGC	102
Db	971	TGGAAGCCTGGTACAGACAGAAATTTTGGGAACCTATTGTACTTTGGGAGATGATAGAC	1030
QY	1022	ACCTCACCAACCGCATCTCTCAGCATGGTTATGCTACCAAGTA	1064
Db	1031	ATCTGACAAACCGAGTGCTCAGCATGGGATATCGACCAAATA	1073
RESULT 15			
US-09-902-939-3			
; Sequence 3, Application US/09902939			
; Publication No. US20030087850A1			
; GENERAL INFORMATION:			
; APPLICANT: Philip DeHaza			
; APPLICANT: Weiliam Chen			
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME			
; FILE REFERENCE: 2055/0H020-US0			
; CURRENT APPLICATION NUMBER: US/09/902,939			
; CURRENT FILING DATE: 2001-07-10			
; NUMBER OF SEQ ID NOS: 6			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 3			
; LENGTH: 1665			
; TYPE: DNA			
; ORGANISM: mouse			
US-09-902-939-3			
Query Match 29.3%; Score 311.6; DB 10; Length 1665;			
Best Local Similarity 60.3%; Pred. No. 2.1e-72;			
Matches 609; Conservative 0; Mismatches 359; Indels 42; Gaps 4;			
QY	61	CGGAGGCTGTGACCATCGCTTCGCGCCTCTCATCTCTGGGCTCATATGACCTCGGCTTAC	120
Db	28	CGTGTGTGGGCACCACTGTGTTTGCCCTGGTAGTGTGGAGGCACTCTGGCGGCTTAT	87
QY	121	GCGCCCGGGGTGCGGTGGCCTCCGATCGCTACGCGCTCTCTGGCCTTCGCGCTTACGGG	180
Db	88	GTGACAGGCTACCAAGTTTATCCACACAGAAAGCACTACCTGTCTCTTGGCCTCTACGGT	147
QY	181	GCCTTCCTTTCAGGCGACCTGGTGGCGCAGAGCCTCTTTCGCGTACTGTGAGCAACCGCGG	240
Db	148	GCCATCTCTGGGTCTACACTCTGCTCATCCAGAGCCTGTGTTGCCCTCTCTGGAGCACCGTCGA	207
QY	241	GTGCGCGCGCGCGCGCG- - - - -GGGCGCGCTGGATGACGACCACCGCGCGCAGTGTGGCG	294
Db	208	ATGCGCAGGCGAGGGCGCCCCCTCAAGCTGCACTGTCTCCAGAGGTGCGGTTCAGTGGCA	267
QY	295	CTGACCACTCTCCGCTACAGAGGAGACCCCGCTACTCTGCGCAGTGCTTCGCGCTCCGCG	354
Db	268	CTCTGCATTGCTGCTTACCAAGAGGACCCCGAATACCTGCGCAAGTGCTTCGCTCAGCT	327
QY	355	CGCGCCCTGTGTATACCGCGCGCGCGCTGCGGCTGCTCATGTGTGTGATGTCACACCGC	414
Db	328	CAGCGCATTCGCTTTTCCAAAC- - - - -CTCAAGGTGGTCACTGGTAGTGGATGCAATCGC	381
QY	415	GCGGAGACCTCTACATGCTGCAGATGTTCCGCGAGTCTTTCGCTCAGCAGGACCCCGCG	474
Db	382	CAGGAATATACCTACATGTTGGACATCTTCCATGAGGTGCTTGGTGGCATGAGCAAGCT	441
QY	475	ACGTACGTGTGGGACGGCAACTACCAACGCCCTTGGGAACCCCGCGCGCGCGCGGTG	534
Db	442	GGCTTCTTTGTGTGGGTAGCAATTTCCA- - - - -TGAGCGGGGTGAAGGA	486
QY	535	GGCGCCGGAGCCTTATCGGGAGGTGGAGGCGGAGGATTCCTGGGCGGCTGGCAGTGGAGGGC	594
Db	487	GAGACAGAGCCAGCCTGACAGGAGGCA TGGAGCGT- - - - -GTGCGAGCT	531
QY	595	CTGGTGAAGACTCGCAGGTGCGTTCGTTGGCGCAGCGCTGGGGCGGCAAGCGCGAGGTC	654
Db	532	GTGGTGTGGGCCAGCACCTTCTCATGCATCATGCAAGTGGGGGGGCAAGCGTGAAGTC	591
QY	655	ATGTACACAGCCTTTCAGGCGCTCGGAGATTTGGTGGACTTACGTGCAAGTCTGTGACTCG	714

Db	592	ATGTACACTGCCCTTCAAGGGCCCTTGGCAACTCAGTGGACTACATCCAGGTGTGTGACTCT	651
Qy	715	GACACAGGTTTGACCCCATGGCACCTCTCGAGCTCGTGGGGTACTTGGACGAGGACCCC	774
Db	652	GACACTGTCTGGACCCGAGCTGCACCAITTGAGATGCTTCGAGTCTTGGGAAGAAGATCCC	711
Qy	775	CGGTTAGGGCTGTGTGGTGGGACGTGCGGATCCTTAAACCCTCTGGACTCTCGGGTCAGC	834
Db	712	CAAGTAGAGGTGTGGAGAGATGTCCAATCTCAAAGAATATGATTCATGGATCTCC	771
Qy	835	TTCCCTAAGCAGCCTGCGATACTGGTAGCCCTCAATGTGGAGCGGGCTTGTCCAGAGCTAC	894
Db	772	TTCTCTGAGCAGTGTGAGGTACTGGATGCGCTTCAAAGTGGAGCGGGCTTGCAGTCCCTAC	831
Qy	895	TTCCACTGTGTATCTCTGCATCAGCGGTCTCTTAGGCGCTATATAGGAATAACTCTTTGCAG	954
Db	832	TTTGGCTGTGTGCAATGTAATTAGTGGGCCCTTTGGGCAATGTAACCGCAACAGCCCTCTTCAG	891
Qy	955	CAGTTTCTTGAGGGCTGGTCAACACGAAAGTTCTCTGGGTACCCACTGTACTTTTGGGGAT	1014
Db	892	CAGTTCTCTGGAGGATTGTTACCATCAGAAGTTCTCTAGGCGCAGCAAGTGCAGCTTTGGGGAT	951
Qy	1015	GACCGGACCTTCCAAACCCGATGCTCAGCATGGGTATGCTACCAAGTA	1064
Db	952	GATTCGGCACCTTACCAACCGAGTCTTGAGTCTTTGGCTACCGGGACTAAGTA	1001

Search completed: March 13, 2005, 17:10:03
Job time : 753.676 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:11:12 ; Search time 5176.05 Seconds
(without alignments)
8399.558 Million cell updates/sec

Title: US-10-672-399-7
Perfect score: 1065
Sequence: 1 atgagacagcaggacgcgc.....tgggtatgtaccacgaagtaa 1065

Scoring table: IDENTITY NUC
Gapop 10_0 , Gapext 1.0

Searched: 45554873 seqs, 20411521753 residues
Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/PCTUS1_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/PCTUS2_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/PCTUS3_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/PCTUS4_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/PCTUS5_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/PCTUS6_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/PCTUS7_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/PCTUS8_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/PCTUS9_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/PCTUS10_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/PCTUS11_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/PCTUS12_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/PCTUS13_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/PCTUS14_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/PCTUS15_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/PCTUS16_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/PCTUS17_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/PCTUS18_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/PCTUS19_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/PCTUS20_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/PCTUS21_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/PCTUS22_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/PCTUS23_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/PCTUS24_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/PCTUS25_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/PCTUS26_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/PCTUS27_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/PCTUS28_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/PCTUS29_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/PCTUS30_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/PCTUS31_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/PCTUS32_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/PCTUS33_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/PCTUS34_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/PCTUS35_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/PCTUS36_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/PCTUS37_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/PCTUS38_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/PCTUS39_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/PCTUS40_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/PCTUS41_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/PCTUS42_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/PCTUS43_COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US099F_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US099G_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US103A_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US103B_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US104A_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US104B_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US105A_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US105B_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US106A_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US107A_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US107B_COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US107C_COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US107D_COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US108A_COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US108B_COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US109A_COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US109B_COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US109C_COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US110_COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
86: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
87: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
88: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
89: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
90: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
91: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
92: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq.*
93: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq.*
94: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
95: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
96: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
97: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
98: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
99: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
100: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
101: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
102: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
103: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
104: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
105: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
106: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
107: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
108: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
109: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
110: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
111: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*
112: /cgn2_6/ptodata/1/pna/US6042_COMB.seq.*
113: /cgn2_6/ptodata/1/pna/US6043_COMB.seq.*
114: /cgn2_6/ptodata/1/pna/US6044_COMB.seq.*
115: /cgn2_6/ptodata/1/pna/US6045_COMB.seq.*
116: /cgn2_6/ptodata/1/pna/US6046_COMB.seq.*

Query Match	99.9%;	Score 1064;	DB 1;	Length 1737;
Best Local Similarity	100.0%;	Pred. No. 2.8e-164;		
Matches 1064;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGAGACAGCAGACGGCGCCCAAGCCCACTCTCTGCAGCCCGCGCTCTCTCGGGCTGGCC	60	
Db	1	ATGAGACAGCAGACGGCGCCCAAGCCCACTCTCTGCAGCCCGCGCTCTCTCGGGCTGGCC	60	
QY	61	CGGAGGGTGCTGACCATCGCTTTCGCCCTGCTCATCTCTGGGCTCATGACTGGGCTTAC	120	
Db	61	CGGAGGGTGCTGACCATCGCTTTCGCCCTGCTCATCTCTGGGCTCATGACTGGGCTTAC	120	
QY	121	GCCGCGGGGGTGCGCTGGGCTCCGATCGCTACGGCCCTCTTGCCCTCTCGGCTCTACGGG	180	
Db	121	GCCGCGGGGGTGCGCTGGGCTCCGATCGCTACGGCCCTCTTGCCCTCTCGGCTCTACGGG	180	
QY	181	GCCTTTCCTTTACGGCACCTGGTGGCGACAGACCTCTTTCGCTACCTTGGAGCACCGGCGG	240	
Db	181	GCCTTTCCTTTACGGCACCTGGTGGCGACAGACCTCTTTCGCTACCTTGGAGCACCGGCGG	240	
QY	241	GTGGCGCGCGCGCGGGGGCGCTGGATGCAGCCACCGCGGCGAGTGTGGGCTGTGACC	300	
Db	241	GTGGCGCGCGCGCGGGGGCGCTGGATGCAGCCACCGCGGCGAGTGTGGGCTGTGACC	300	
QY	301	ATCTCGGCTTACAGGAGGACCCCGGCTACCTGCGCAGTGCTTGGCGTCCGCCCGCGCC	360	

```

RESULT 3
US-10-672-399-1
; Sequence 1, Application US/10672399
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-1

```

Query Match 99.9%; Score 1064; DB 58; Length 1737;
Best Local Similarity 100.0%; Pred. No. 2.8e-164;
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Query Match 99.9%; Score 1064; DB 52; Length 2088;
Best Local Similarity 100.0%; Pred. No. 2.8e-164;
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGACAGCAGGAGCGCGCCCAAGCCCACTCTCTGAGCCCGCGCTGCTCCGGGCTGGCC 60
DB |||||
QY 36 ATGAGACAGCAGGAGCGCGCCCAAGCCCACTCTCTGAGCCCGCGCTGCTCCGGGCTGGCC 95
DB |||||
QY 61 CGGAGGGTGTGACCATCGCTTTCGCCCTGTCTATCTCTGGCCCTCATGACCTGGGCTTAC 120
DB |||||
QY 96 CGGAGGGTGTGACCATCGCTTTCGCCCTGTCTATCTCTGGCCCTCATGACCTGGGCTTAC 155
DB |||||
QY 121 GCCGCGGGGTGCGGCTCGATCGCTACCGGCTCTCTGGCCCTTGGGCTTACCGG 180
DB |||||
QY 156 GCCGCGGGGTGCGGCTCGATCGCTACCGGCTCTCTGGCCCTTGGGCTTACCGG 215
DB |||||
QY 181 GCCTTCTCTTTCAGCGCACCTGCTGGCGCAGAGCTCTCTGGGCTACCTGGAGCACCGGCGG 240
DB |||||
QY 216 GCCTTCTCTTTCAGCGCACCTGCTGGCGCAGAGCTCTCTGGGCTACCTGGAGCACCGGCGG 275
DB |||||
QY 241 GTGCGCGCGCGCGCGCGGCGCGCTGTGATGACGACACCGCGCGAGTGTGGGCTGAC 300
DB |||||
QY 276 GTGCGCGCGCGCGCGCGGCGCGCTGTGATGACGACACCGCGCGAGTGTGGGCTGAC 335
DB |||||
QY 301 ATCTCCGCTTACAGGAGGACCGCGCTTCTCTGGCGAGTCTCTGGGCTCGCGCGCGCC 360
DB |||||
QY 336 ATCTCCGCTTACAGGAGGACCGCGCTTCTCTGGCGAGTCTCTGGGCTCGCGCGCGCC 395
DB |||||
QY 361 CTGCTGTACCCGCGCGCGCGCTGCGCTCTCTATGTTGGTGGATGGCAACCGGCGCGAG 420
DB |||||
QY 396 CTGCTGTACCCGCGCGCGCGCTGCGCTCTCTATGTTGGTGGATGGCAACCGGCGCGAG 455
DB |||||
QY 421 GACCTCTACATGTGTGACATGTTCCGCGAGGTCCTCGCTGACGAGGACCCCGCGCACG 480
DB |||||
QY 456 GACCTCTACATGTGTGACATGTTCCGCGAGGTCCTCGCTGACGAGGACCCCGCGCACG 515
DB |||||
QY 481 GTGTGGGACCGCACTACCAAGGACCGCGCTTCTCTGGCGAGTCTCTGGGCTCGCGCGCC 540
DB |||||
QY 516 GTGTGGGACCGCACTACCAAGGACCGCGCTTCTCTGGCGAGTCTCTGGGCTCGCGCGCC 575
DB |||||
QY 541 GGAGCTATCCGAGGTGGAGCGGAGGATCTCTGGCGGCTGGAGTGGAGCGCTGGTG 600
DB |||||
QY 576 GGAGCTATCCGAGGTGGAGCGGAGGATCTCTGGCGGCTGGAGTGGAGCGCTGGTG 635
DB |||||
QY 601 AGGACTCGAGGTGCTGTGCGTGGCGCAGCGCTGGGCGCGCAAGCGCGAGGTCA 660
DB |||||
QY 636 AGGACTCGAGGTGCTGTGCGTGGCGCAGCGCTGGGCGCGCAAGCGCGAGGTCA 695
DB |||||
QY 661 ACAGCTTCAAGGCGCTCGGAGATTCGGTGGACTACGTCGAGGTCGTGACTCGGACACA 720
DB |||||
QY 696 ACAGCTTCAAGGCGCTCGGAGATTCGGTGGACTACGTCGAGGTCGTGACTCGGACACA 755
DB |||||
QY 721 AGGTGGACCCATGGCACTGCTGAGCTCTGGGCTTCTGGGCTTCTGGAGGACCCCGGGA 780
DB |||||
QY 756 AGTTGGACCCATGGCACTGCTGAGCTCTGGGCTTCTGGGCTTCTGGAGGACCCCGGGA 815
DB |||||
QY 781 GGGGCTGTGTGGGACGTCGAGTCTTAACTCTGGACTCTGGGCTCAGCTTCTTA 840
DB |||||
QY 816 GGGGCTGTGTGGGACGTCGAGTCTTAACTCTGGACTCTGGGCTCAGCTTCTTA 875
DB |||||
QY 841 AGCAGCTGCATACTGGGTAGCTTCAATGTGGAGCGGCTTGTTCAGAGCTACTTCCAC 900
DB |||||
QY 876 AGCAGCTGCATACTGGGTAGCTTCAATGTGGAGCGGCTTGTTCAGAGCTACTTCCAC 935
DB |||||
QY 901 TGTGTATCTGCATCAGCGGCTCTTAGGCTATATAGGAATACTCTTGCAGCAGTTT 960
DB |||||
QY 936 TGTGTATCTGCATCAGCGGCTCTTAGGCTATATAGGAATACTCTTGCAGCAGTTT 995
DB |||||
QY 961 CTTGAGGCTTGGTACAAACAGAGTTCTTGGGTACCCACTGTATCTTTTGGGATGA 1020
DB |||||
QY 996 CTTGAGGCTTGGTACAAACAGAGTTCTTGGGTACCCACTGTATCTTTTGGGATGA 1055
DB |||||

QY 1021 CACCTCACCAACCGCATGCTCAGCATGGGTTATGCTACCAAGTA 1064
DB |||||
QY 1056 CACCTCACCAACCGCATGCTCAGCATGGGTTATGCTACCAAGTA 1099
DB |||||

RESULT 7
US-60-257-537-222
; Sequence 222, Application US/60257537
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen
; APPLICANT: Schafer, Alan
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH OSTEOARTHRITIS
; FILE REFERENCE: GX-0022 P
; CURRENT APPLICATION NUMBER: US/60/257,537
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 499
; SOFTWARE: PERL Program
; SEQ ID NO 222
; LENGTH: 2088
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: GB:NM_001523
US-60-257-537-222

Query Match 99.9%; Score 1064; DB 95; Length 2088;
Best Local Similarity 100.0%; Pred. No. 2.8e-164;
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGACAGCAGGAGCGCGCCCAAGCCCACTCTCTGAGCCCGCGCTGCTCCGGGCTGGCC 60
DB |||||
QY 36 ATGAGACAGCAGGAGCGCGCCCAAGCCCACTCTCTGAGCCCGCGCTGCTCCGGGCTGGCC 95
DB |||||
QY 61 CGGAGGGTGTGACCATCGCTTTCGCCCTGTCTATCTCTGGGCTTACGACCTGGGCTTAC 120
DB |||||
QY 96 CGGAGGGTGTGACCATCGCTTTCGCCCTGTCTATCTCTGGGCTTACGACCTGGGCTTAC 155
DB |||||
QY 121 GCCGCGGGGTGCGGCTCGATCGCTTCCGATCGCTTCCGATCGCTTCTGGGCTTCTGGGCTTAC 180
DB |||||
QY 156 GCCGCGGGGTGCGGCTCGATCGCTTCCGATCGCTTCTGGGCTTCTGGGCTTCTGGGCTTAC 215
DB |||||
QY 181 GCCTTCTCTTTCAGCGCACCTGCTGGCGCAGAGCTCTTTCGCTGACCTGGAGCACCGGCGG 240
DB |||||
QY 216 GCCTTCTCTTTCAGCGCACCTGCTGGCGCAGAGCTCTTTCGCTGACCTGGAGCACCGGCGG 275
DB |||||
QY 241 GTGCGCGCGCGCGCGCGGCGCGCTGTGATGACGACACCGCGCGAGTGTGGGCTGAC 300
DB |||||
QY 276 GTGCGCGCGCGCGCGGCGCGCTGTGATGACGACACCGCGCGAGTGTGGGCTGAC 335
DB |||||
QY 301 ATCTCCGCTTACAGGAGGACCGCGCTTCTCTGGCGAGTCTCTGGGCTCGCGCGCGCC 360
DB |||||
QY 336 ATCTCCGCTTACAGGAGGACCGCGCTTCTCTGGCGAGTCTCTGGGCTCGCGCGCGCC 395
DB |||||
QY 361 CTGCTGTACCCGCGCGCGCTGCGCTCTCTATGTTGGTGGATGGCAACCGGCGCGAG 420
DB |||||
QY 396 CTGCTGTACCCGCGCGCGCGCTGCGCTCTCTATGTTGGTGGATGGCAACCGGCGCGAG 455
DB |||||
QY 421 GACCTCTACATGTGTGACATGTTCCGCGAGGTCCTCGCTGACGAGGACCCCGCGCACG 480
DB |||||
QY 456 GACCTCTACATGTGTGACATGTTCCGCGAGGTCCTCGCTGACGAGGACCCCGCGCACG 515
DB |||||
QY 481 GTGTGGGACCGCACTACCAAGGACCGCGCTTCTCTGGCGAGTCTCTGGGCTCGCGCGCC 540
DB |||||
QY 516 GTGTGGGACCGCACTACCAAGGACCGCGCTTCTCTGGCGAGTCTCTGGGCTCGCGCGCC 575
DB |||||
QY 541 GGAGCTATCCGAGGTGGAGCGGAGGATCTCTGGCGGCTGGAGTGGAGCGCTGGTG 600
DB |||||
QY 576 GGAGCTATCCGAGGTGGAGCGGAGGATCTCTGGCGGCTGGAGTGGAGCGCTGGTG 635
DB |||||
QY 601 AGGACTCGAGGTGCTGTGCGTGGCGCAGCGCTGGGCGCGCAAGCGCGAGGTCA 660
DB |||||
QY 636 AGGACTCGAGGTGCTGTGCGTGGCGCAGCGCTGGGCGCGCAAGCGCGAGGTCA 695
DB |||||

661 ACAGCCTTCAAGCGCTCGAGATTCGGTGGACTACGTGACGTCTGTGACTCGGACACA 720
696 ACAGCCTTCAAGCGCTCGAGATTCGGTGGACTACGTGACGTCTGTGACTCGGACACA 755
721 AGGTTGGACCCCATGCACTGCTGGAGCTGCTGGGGTACTGACAGAGACCCCGGGTA 780
756 AGGTTGGACCCCATGCACTGCTGGAGCTGCTGGGGTACTGACAGAGACCCCGGGTA 815
781 GGGGCTGTTGGTGGGACGTCGGATCCTTAACCTCTGAGCTCCTGGGTTCAGTTCCTTA 840
816 GGGGCTGTTGGTGGGACGTCGGATCCTTAACCTCTGAGCTCCTGGGTTCAGTTCCTTA 875
841 AGCAGCTCGGATACTGGGTAGCTCTCAATGTGGAGCGGCTTGTGACAGACTACTTCCAC 900
876 AGCAGCTCGGATACTGGGTAGCTCTCAATGTGGAGCGGCTTGTGACAGACTACTTCCAC 935
901 TGTGTATCTGTCATCAGCGGCTCTAGGCTATATAGGAATACCTCTTGCAGGAGTTT 960
936 TGTGTATCTGTCATCAGCGGCTCTAGGCTATATAGGAATACCTCTTGCAGGAGTTT 995
961 CTTGAGGCTGGTACAAACAGAGTTTCTGGGTACCCACTGTACTTTTGGGGATGACCG 1020
996 CTTGAGGCTGGTACAAACAGAGTTTCTGGGTACCCACTGTACTTTTGGGGATGACCG 1055
1021 CACCTCACCAACCGCATGCTCAGCATGGGTATGTACCAAGTA 1064
1056 CACCTCACCAACCGCATGCTCAGCATGGGTATGTACCAAGTA 1099

RESULT 8
US-60-278-258-16271
; Sequence 16271, Application US/60278258
; GENERAL INFORMATION:
; APPLICANT: Morris, Macdonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; TITLE OF INVENTION: Polymorphisms Identified Thereby
; FILE REFERENCE: GX-0010-1 P
; CURRENT APPLICATION NUMBER: US/60/278,258
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 17730
; SOFTWARE: PERL Program
; SEQ ID NO 16271
; LENGTH: 2098
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 335023.1
US-60-278-258-16271

Query Match 99.9%; Score 1064; DB 97; Length 2098;
Best Local Similarity 100.0%; Pred. No. 2.8e-164;
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAGACAGCAGCAGCGCCGAGCCCACTCTCTGAGCCCGCGGCTGCTCGGCTCGGC 60
48 ATGAGACAGCAGCAGCGCCGAGCCCACTCTCTGAGCCCGCGGCTGCTCGGCTCGGC 107
61 CGAGGGTGTGACCATCGCTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
108 CGAGGGTGTGACCATCGCTTCGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 167
121 GCCCGCGGGTCCGCTGCTCCGATCGCTACGGCTCTCTGGCTTGGGCTCTACGGG 180
168 GCCCGCGGGTCCGCTGCTCCGATCGCTACGGCTCTCTGGCTTGGGCTCTACGGG 227
181 GCCTTCCTTTAGCGGCACTGTGGGCGAGAGCTCTTGGGCTACCTGGAGCAGCGGG 240
228 GCCTTCCTTTAGCGGCACTGTGGGCGAGAGCTCTTGGGCTACCTGGAGCAGCGGG 287

241 GTGGGGCGCGCGCGCGCGCGCTGGAATGACGACCCGCGCAGTGTGGCGCTGACC 300
288 GTGGGGCGCGCGCGCGCGCGCTGGAATGACGACCCGCGCAGTGTGGCGCTGACC 347
301 ATCTCCGCTTACAGAGAGACCCCGGCTACCTGCGCACTGCTGCGCTCGCGCCGCGC 360
348 ATCTCCGCTTACAGAGAGACCCCGGCTACCTGCGCACTGCTGCGCTCGCGCCGCGC 407
361 CTGCTGTATCCCGCGCGCGCGCTGCGCTCCTCATGTTGGTGGTGGTGGTGGTGGTGG 420
408 CTGCTGTATCCCGCGCGCGCGCTGCGCTCCTCATGTTGGTGGTGGTGGTGGTGGTGG 467
421 GACCTCTTACATGGTGCAGCATGTTCCGCGCAGAGTCTTCTGCTGACGAGAGACCCCGC 480
468 GACCTCTTACATGGTGCAGCATGTTCCGCGCAGAGTCTTCTGCTGACGAGAGACCCCGC 527
481 GTGTGGAGCGCACTACCAACAGAGCTTGGGAAACCCGCGCGCGCGCGCGCTGGCGCC 540
528 GTGTGGAGCGCACTACCAACAGAGCTTGGGAAACCCGCGCGCGCGCGCTGGCGCC 587
541 GGAGCTATCCGAGAGTGGAGCGGAGGATCCTGGCGGCTGGCAGTGGAGCGCTGGTG 600
588 GGAGCTATCCGAGAGTGGAGCGGAGGATCCTGGCGGCTGGCAGTGGAGCGCTGGTG 647
601 AGGACTCGCAGGTGCTGTGCTGCGCAGCGCTGGGGCGCAAGCGCGAGGTCATGTAC 660
648 AGGACTCGCAGGTGCTGTGCTGCGCAGCGCTGGGGCGCAAGCGCGAGGTCATGTAC 707
661 ACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGACGAGTCTGTGACTCGGACACA 720
708 ACAGCCTTCAAGCGCTCGGAGATTCGGTGGACTACGTGACGAGTCTGTGACTCGGACACA 767
721 AGGTTGGACCCCATGCACTGCTGAGCTCTGCGGGTACTGACGAGAGACCCCGGGTA 780
768 AGGTTGGACCCCATGCACTGCTGAGCTCTGCGGGTACTGACGAGAGACCCCGGGTA 827
781 GGGGCTGTTGGTGGGAGCTGCGGATCCTTAACCTCTGGAATCTCTGGGTTCAGTTCCTTA 840
828 GGGGCTGTTGGTGGGAGCTGCGGATCCTTAACCTCTGGAATCTCTGGGTTCAGTTCCTTA 887
841 AGCAGCTCGGATACCTGAGTACCTCAATGTGGAGCGGCTTGTGAGAGTACTTCCAC 900
888 AGCAGCTCGGATACCTGAGTACCTCAATGTGGAGCGGCTTGTGAGAGTACTTCCAC 947
901 TGTGTATCTGCACTCAGCGGCTCTTAGGCTATATAGGAATAAACCCTCTTGCAGAGTTT 960
948 TGTGTATCTGCACTCAGCGGCTCTTAGGCTATATAGGAATAAACCCTCTTGCAGAGTTT 1007
961 CTTGAGGCTGGTACAAACAGAGTTCCTGGGTACCCACTGCTGAGTATGCGGATGACCGG 1020
1008 CTTGAGGCTGGTACAAACAGAGTTCCTGGGTACCCACTGCTGAGTATGCGGATGACCGG 1067
1021 CACCTCACCAACCGCATGCTCAGCATGGGTATGTACCAAGTA 1064
1068 CACCTCACCAACCGCATGCTCAGCATGGGTATGTACCAAGTA 1111

RESULT 9
US-10-042-523-1
; Sequence 1, Application US/10042523
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/042,523
FILING DATE: 19-Oct-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/635,552A
FILING DATE: 22-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: LK995-07
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2116 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 36..1769
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-042-523-1

Query Match 99.9%; Score 1064; DB 46; Length 2116;
Best Local Similarity 100.0%; Pred. No. 2.8e-164;
Matches 1064; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGACAGCAGGAGCGCCAGCCCACTCTGACGCGCGCGCTGCTCGGCGCTGGCC 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 36 ATGACAGCAGGAGCGCCAGCCCACTCTGACGCGCGCGCTGCTCGGCGCTGGCC 95
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 CGGAGGGTGCTGACCATCGCTTCCCTTGCTCATCTCTGGGCTCATGACCTGGGCGTAC 120
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 96 CGGAGGGTGCTGACCATCGCTTCCCTTGCTCATCTCTGGGCTCATGACCTGGGCGTAC 155
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 GCGCGCGGGTGCGGCTGCTCCGATCGCTACGCGCTCTGGGCTCTGGGCTCTACGGG 180
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 156 GCGCGCGGGTGCGGCTGCTCCGATCGCTACGCGCTCTGGGCTCTGGGCTCTACGGG 215
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 181 GCCTTCTTTCAGCGCACCTGCTGCGCAGAGCCTCTTTCGGGTACTCTGGAGCACCGCGG 240
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 216 GCCTTCTTTCAGCGCACCTGCTGCGCAGAGCCTCTTTCGGGTACTCTGGAGCACCGCGG 275
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 241 GTGGCGGCGGCGCGCGCGCGCTGGATGAGCCACCGCGCGCAGTGTGGCGCTGACC 300
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 276 GTGGCGGCGGCGCGCGCGCGCTGGATGAGCCACCGCGCGCAGTGTGGCGCTGACC 335
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 301 ATCTCCGCTTACAGAGAGACCCCGGCTGCTGCGCCAGTGCTGGCGTCCGCGCGCC 360
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 336 ATCTCCGCTTACAGAGAGACCCCGGCTGCTGCGCCAGTGCTGGCGTCCGCGCGCC 395
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 361 CTGCTGTACCGCGCGCGCGCTGCGCTGCTCATGCTGCTGAGTGGCAACCGCGCGCAG 420
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 396 CTGCTGTACCGCGCGCGCGCTGCGCTGCTCATGCTGCTGAGTGGCAACCGCGCGCAG 455
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 421 GACCTCTACATGGTGCATGTTCCGCGAGGTCTTTCGCTGACGAGAGACCCCGCCACGTCAC 480
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 456 GACCTCTACATGGTGCATGTTCCGCGAGGTCTTTCGCTGACGAGAGACCCCGCCACGTCAC 515
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 481 GTGTGGGAGCGGAATACACAGCCCTGGGAAACCCCGCGCGCGCGGCGCGGTGGGCGCC 540
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 516 GTGTGGGAGCGGAATACACAGCCCTGGGAAACCCCGCGCGCGCGGCGCGGTGGGCGCC 575
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 541 GGAGCCTATCGGAGGTGAGCGCGAGGATCTCTGGGCGGCTGGCAGTGGAGCGCGTGTG 600

DB 576 GGAGCCTATCGGAGGTGAGCGCGAGGATCTTGGCGCGCTGGCAGTGGAGCGCTGGTG 635
QY 601 AGGACTCCAGGTGCGTGTGCGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCAATGTAC 660
DB 636 AGGACTCGAGGTGCGTGTGCGCGCAGCGCTGGGGCGGCAAGCGCGAGGTCAATGTAC 695
QY 661 ACAGCCTTCAAGGCGCTCGGAGATTCCGTGACACTACGTGACAGGTCTGTGACCTCGGACACA 720
DB 696 ACAGCCTTCAAGGCGCTCGGAGATTCCGTGACACTACGTGACAGGTCTGTGACCTCGGACACA 755
QY 721 AGGTTGGACCCCATGGCACTCTGAGACTCTGGGGGTACTGGACGAGGACCCCGGGTA 780
DB 756 AGGTTGGACCCCATGGCACTCTGAGACTCTGGGGGTACTGGACGAGGACCCCGGGTA 815
QY 781 GGGGCTGTGTGGGAGCGTCCGATCCTTAACCTCTGGACTCTCTGGGTGAGTTCCTTA 840
DB 816 GGGGCTGTGTGGGAGCGTCCGATCCTTAACCTCTGGACTCTCTGGGTGAGTTCCTTA 875
QY 841 AGCAGCCTCGATACTGGGTAGCCTTCAATGAGCGGCGCTTGTGACAGCTACTTCCAC 900
DB 876 AGCAGCCTCGATACTGGGTAGCCTTCAATGAGCGGCGCTTGTGACAGCTACTTCCAC 935
QY 901 TGTGTATCTGCATCAGCGGTCTCTAGGCTTATAGGAATAACCTCTTGACGAGTGT 960
DB 936 TGTGTATCTGCATCAGCGGTCTCTAGGCTTATAGGAATAACCTCTTGACGAGTGT 995
QY 961 CTTGAGGCGTGGTAAACAGAGATTCTCTGGGTACCCACTGCTATTTGGGATGACCGG 1020
DB 996 CTTGAGGCGTGGTAAACAGAGATTCTCTGGGTACCCACTGCTATTTGGGATGACCGG 1055
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGTTATGCTACCAAGTA 1064
DB 1056 CACCTCACCAACCGCATGCTCAGCATGGTTATGCTACCAAGTA 1099

RESULT 10

PCT-US02-41225A-224
; GENERAL INFORMATION:
; SEQUENCE 224, Application PC/TUS0241225A
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: JONES, Karen Anne
; APPLICANT: SCHAFER, Alan
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH OSTEOARTHRITIS
; FILE REFERENCE: GR-0022 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/41225A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/342,603
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 499
; SOFTWARE: PERL Program
; SEQ ID NO 224
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: HASI_cds.1
PCT-US02-41225A-224

Query Match 99.8%; Score 1062.4; DB 1; Length 1737;
Best Local Similarity 99.9%; Pred. No. 5.2e-164;
Matches 1063; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACAGCAGGAGCGGCGCCCAAGCCCACTCTGACGCGCGCGCTGCTCGGCGCTGGCC 60
DB 1 ATGACAGCAGGAGCGGCGCCCAAGCCCACTCTGACGCGCTGCTCGGCGCTGGCC 60
QY 61 CGGAGGGTGCTGACCATCGCTTCCGCTGCTCATCTCTGGGCTCATGACCTGGGCGCTAC 120
DB 61 CGGAGGGTGCTGACCATCGCTTCCGCTGCTCATCTCTGGGCTCATGACCTGGGCGCTAC 120
QY 121 GCGCGCGGGTGCGGCTGGGCTCGGCTCGATCGCTTCTGGGCTTCTCGGCTCTACGGG 180

Db 121 GCCCGCGGGTCCCGTGGCTCCATCGCTACGCGCTCTCGGCTCTACGGG 180
Qy 181 GCCTTCCTTTAGCGCACCTGGTGGCGAGAGCTTCTGGCTACTGGAGCACGGCGG 240
Db 181 GCCTTCCTTTAGCGCACCTGGTGGCGAGAGCTTCTGGCTACTGGAGCACGGCGG 240
Qy 241 GTGGCGGGCGGCGCGGGGCGCTGTGATGACGACCGCGCGAGTGTGGCGGTACC 300
Db 241 GTGGCGGGCGGCGCGGGGCGCTGTGATGACGACCGCGCGAGTGTGGCGGTACC 300
Qy 301 ATCTCGGCTACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 301 ATCTCGGCTACACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Qy 361 CTGCTGTATACCGCGCGGCGCTGGCGCTCTCATGTGGTGGTGGTGGTGGTGGT 420
Db 361 CTGCTGTATACCGCGCGGCGCTGGCGCTCTCATGTGGTGGTGGTGGTGGTGGT 420
Qy 421 GACCTCTACATGGTTCGATGTTTCGCGAGAGTTCCTGCTGAGAGAGAGAGAGAG 480
Db 421 GACCTCTACATGGTTCGCGAGAGTTCCTGCTGAGAGAGAGAGAGAGAGAGAGAG 480
Qy 481 GTGTGGGAGCGCAATACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 GTGTGGGAGCGCAATACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy 541 GGAGGCTATCGGAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 541 GGAGGCTATCGGAGAGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Qy 601 AGGACTCGAGGTGCTGTGCGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 601 AGGACTCGAGGTGCTGTGCGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy 661 ACAGGCTTCAGGCGCTCGAGATTCGCTGGAGTTCGCTGGAGTTCGCTGGAGTTC 720
Db 661 ACAGGCTTCAGGCGCTCGAGATTCGCTGGAGTTCGCTGGAGTTCGCTGGAGTTC 720
Qy 721 AGTTGGAG 780
Db 721 AGTTGGAG 780
Qy 781 GGGGCTGTGTGGGAGAGTGGGATCTTAACTCTGGGCTGGGCTGGGCTGGGCTGGG 840
Db 781 GGGGCTGTGTGGGAGAGTGGGATCTTAACTCTGGGCTGGGCTGGGCTGGGCTGGG 840
Qy 841 AGCAGCTGCGATACCTGGGTAGCTTCAATGTGGAGAGAGAGAGAGAGAGAGAGAG 900
Db 841 AGCAGCTGCGATACCTGGGTAGCTTCAATGTGGAGAGAGAGAGAGAGAGAGAGAG 900
Qy 901 TGTGTATCTGATCAGGAGTCTCTAGGCTTATAGGAATACTCTGGAGAGAGAGAG 960
Db 901 TGTGTATCTGATCAGGAGTCTCTAGGCTTATAGGAATACTCTGGAGAGAGAGAG 960
Qy 961 CTGAGGCTGTGTACCAACAGAGTTCCTGGGTACCCACTGTACTTTTGGGATGACCGG 1020
Db 961 CTGAGGCTGTGTACCAACAGAGTTCCTGGGTACCCACTGTACTTTTGGGATGACCGG 1020
Qy 1021 CACCTACCAACCGGATGCTAGCATGGGTATGTACCAAGTA 1064
Db 1021 CACCTACCAACCGGATGCTAGCATGGGTATGTACCAAGTA 1064

RESULT 11
US-60-257-537-224
; Sequence 224, Application US/60257537
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen
; APPLICANT: Schaefer, Alan
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH OSTEOARTHRITIS
; FILE REFERENCE: GX-0022 P
; CURRENT APPLICATION NUMBER: US/60/257,537

; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 499
; SOFTWARE: PERL Program
; SEQ ID NO 224
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: HAS1_cds.1
US-60-257-537-224

Query Match 99.8%; Score 1062.4; DB 95; Length 1737;
Best Local Similarity 99.9%; Pred. No. 5.2e-164;
Matches 1063; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGAGACAGCAGGAG 60
Db 1 ATGAGACAGCAGGAG 60
Qy 61 CGGAGGGTGTGACCATCGCCTTCCGCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
Db 61 CGGAGGGTGTGACCATCGCCTTCCGCTGCTCATCTCTGGGCTCATGACCTGGGCTTAC 120
Qy 121 GCGCGCGGGGTGCGCGCTGCGCTTCCGATCGCTACGCGCTTCTGGGCTTCTGGGCTTAC 180
Db 121 GCGCGCGGGGTGCGCGCTGCGCTTCCGATCGCTACGCGCTTCTGGGCTTCTGGGCTTAC 180
Qy 181 GCCTTCCTTTTTCAGCGCACCTGGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 181 GCCTTCCTTTTTCAGCGCACCTGGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Qy 241 GTGGCGGCGGCGGCGGCGGCGGCGGCTGGATTCAGCAGCAGCGCGAGTGTGGCGCTGACC 300
Db 241 GTGGCGGCGGCGGCGGCGGCGGCGGCTGGATTCAGCAGCAGCGCGAGTGTGGCGCTGACC 300
Qy 301 ATCTCGGCTTACAG 360
Db 301 ATCTCGGCTTACAG 360
Qy 361 CTGCTGTATACCGCGCGGCGGCGGCTGGGCTCTCATGTGGTGGTGGTGGTGGTGGTGG 420
Db 361 CTGCTGTATACCGCGCGGCGGCGGCTGGGCTCTCATGTGGTGGTGGTGGTGGTGGTGG 420
Qy 421 GACCTCTACATGGTTCGACATGTTCCGCGAGTCTTCTGCTGACGAGAGAGAGAGAGAG 480
Db 421 GACCTCTACATGGTTCGACATGTTCCGCGAGTCTTCTGCTGACGAGAGAGAGAGAGAG 480
Qy 481 GTGTGGGAGCGCAATACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 GTGTGGGAGCGCAATACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Qy 541 GGAGGCTATCGGAGAGTGGAG 600
Db 541 GGAGGCTATCGGAGAGTGGAG 600
Qy 601 AGGACTCGAGGTGCTGTGCGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 601 AGGACTCGAGGTGCTGTGCGTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Qy 661 ACAGGCTTCAGGCGCTCGAGATTCGCTGGAGTTCGCTGGAGTTCGCTGGAGTTC 720
Db 661 ACAGGCTTCAGGCGCTCGAGATTCGCTGGAGTTCGCTGGAGTTCGCTGGAGTTC 720
Qy 721 AGTTGGAG 780
Db 721 AGTTGGAG 780
Qy 781 GGGGCTGTGTGGGAGAGTGGGATCTTAACTCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 840
Db 781 GGGGCTGTGTGGGAGAGTGGGATCTTAACTCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 840
Qy 841 AGCAGCTGCGATACCTGGGTAGCTTCAATGTGGAGAGAGAGAGAGAGAGAGAGAGAG 900
Db 841 AGCAGCTGCGATACCTGGGTAGCTTCAATGTGGAGAGAGAGAGAGAGAGAGAGAGAG 900

Db	841	AGCAGCCTCGCATACGGTAGCCTTCAAATGTGGAGCGGGCTTGTCCAGAGCTACTTCCAC	900
Qy	901	TGTTGTAATCTTCGCATCAGGGTCCTCTAGGCCATATATAGGAATAACCTCTTCGACGAGTTT	960
Db	901	TGTTGTAATCTTCGCATCAGGGTCCTCTAGGCCATATATAGGAATAACCTCTTCGACGAGTTT	960
Qy	961	CTTGAGCGCTGGTACAAACCAGAAAGTTCTGGGTACCCACTGTACTTTTGGGGATGACCGG	1020
Db	961	CTTGAGCGCTGGTACAAACCAGAAAGTTCTGGGTACCCACTGTACTTTTGGGGATGACCGG	1020
Qy	1021	CACCTCACCAACCGCATGCTCAGCATGGGTATGCTACCAAGTA	1064
Db	1021	CACCTCACCAACCGCATGCTCAGCATGGGTATGCTACCAAGTA	1064

```

RESULT 12
PCT-US02-41225A-226
; Sequence 226, Application PC/TUS0241225A
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: JONES, Karen Anne
; APPLICANT: SCHAPER, Alan
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH OSTEOARTHRITIS
; FILE REFERENCE: GX-0022 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/41225A
; CURRENT FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: US 60/342,603
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 499
; SOFTWARE: PERL Program
; SEQ ID NO 226
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: HAS1_mrna_build.1
PCT-US02-41225A-226

```

[illegible]

```

RESULT 13
US-60-257-537-226
; Sequence 226, Application US/60257537
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen
; APPLICANT: Schafer, Alan
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH OSTEOARTHRITIS
; FILE REFERENCE: GX-0022 P
; CURRENT APPLICATION NUMBER: US/60/257,537
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 499
; SOFTWARE: PERL Program
; SEQ ID NO 226
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: HAS1_mrna_build.1
US-60-257-537-226

```

	Query Match	99.8%	Score 1062.4;	DB 95;	Length 2087;
	Best Local Similarity	99.9%	Pred. No. 5.1e-164;		
	Matches 1063; Conservative	0;	Mismatches -1;	Indels 0;	Gaps 0
Qy	1	ATGAGACAGCAGGAGCGGCCCAAGCCCACTCCTGCAGCCCGCGCGTGTTCGGGCTGGCC	60		
Dd	36	ATGAGACAGCAGGAGCGGCCCAAGCCCACTCCTGCAGCTCGCGTGTTCGGGCTGGCC	95		

QY 841 AGCAGCTGCGATATCTGGGTAGCTTCAATGTGGAGCGGGTGTGTCAGAGCTACTTCCAC 900
Db |||||
QY 876 AGCAGCTGCGATATCTGGGTAGCTTCAATGTGGAGCGGGTGTGTCAGAGCTACTTCCAC 935
Db |||||
QY 901 TGTGTATCTGTCATCAGCGGTCTCTAGGCTTATATAGGAATAAATCTTTCAGAGCTTT 960
Db |||||
QY 936 TGTGTATCTGTCATCAGCGGTCTCTAGGCTTATATAGGAATAAATCTTTCAGAGCTTT 995
Db |||||
QY 961 CTTGAGGCTGTGTACAAACAGAGTCTCTGGGTACCACTGTCATTTTGGGGATGACCGG 1020
Db |||||
QY 996 CTTGAGGCTGTGTACAAACAGAGTCTCTGGGTACCACTGTCATTTTGGGGATGACCGG 1055
Db |||||
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGTATGCTACCAAGTA 1064
Db |||||
QY 1056 CACCTCACCAACCGCATGCTCAGCATGGTATGCTACCAAGTA 1099
Db |||||
RESULT 15
US-10-940-774-5659
; Sequence 5659, Application US/10940774
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/10/940,774
; CURRENT FILING DATE: 2004-09-15
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5659
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Human
US-10-940-774-5659

Query Match 99.6%; Score 1060.8; DB 66; Length 2087;
Best Local Similarity 99.8%; Pred. No. 9.3e-164;
Matches 1062; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGAGACAGCAGGACGCGCCCAAGCCCACTCTCTGAGCCCGCGCTGTCTCGGCTGCGC 60
Db |||||
QY 61 CGGAGGTGTGACCATGCGCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 120
Db |||||
QY 96 CGGAGGTGTGACCATGCGCTTGGCCCTGCTCATCTGGGCTCATGACCTGGGCTTAC 155
Db |||||
QY 121 GCGCGCGGGTGGCGCTGGCTTCCGATCGCTACGCGCTTCTGGGCTTCTACGGG 180
Db |||||
QY 156 GCGCGCGGGTGGCGCTGGCTTCCGATCGCTACGCGCTTCTGGGCTTCTACGGG 215
Db |||||
QY 181 GCCTTCTTTTTCAGCGCACCTGGTGGCGCAGAGCTTTCGGCTACTTGGAGCACCGGCGG 240
Db |||||
QY 216 GCCTTCTTTTTCAGCGCACCTGGTGGCGCAGAGCTTTCGGCTACTTGGAGCACCGGCGG 275
Db |||||
QY 241 GTGGGCGGGCGGGCGGGCGGCTGGATGACGAGCACCGGCGGCGAGTGGCGCTGACC 300
Db |||||
QY 276 GTGGGCGGGCGGGCGGGCGGCTGGATGACGAGCACCGGCGGCGAGTGGCGCTGACC 335
Db |||||
QY 301 ATCTCCGCTACACAGGAGACCGCGCTACCTGGCGAGTGCCTGGCGCTCCGCGCGGCGC 360
Db |||||
QY 336 ATCTCCGCTACACAGGAGACCGCGCTACCTGGCGAGTGCCTGGCGCTCCGCGCGGCGC 395
Db |||||
QY 361 CTGCTGTATACCGCGCGCGCTGCGCTCTCATGTGTGGATGGCAACCGCGCGGAG 420
Db |||||
QY 396 CTGCTGTATACCGCGCGCGCTGCGCTCTCATGTGTGGATGGCAACCGCGCGGAG 455
Db |||||

Search completed: March 13, 2005, 16:17:29
Job time : 5178.05 secs

QY 421 GACCTCTACATGTGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGTCAC 480
Db |||||
QY 456 GACCTCTACATGTGTCGACATGTTCCGCGAGGTCTTCGCTGACGAGGACCCCGCCACGTCAC 515
Db |||||
QY 481 GTGTGGGACGGCAACTACCAACAGCCCTTGGGAACCCCGCGCGCGCGCGGTGGCGGCC 540
Db |||||
QY 516 GTGTGGGACGGCAACTACCAACAGCCCTTGGGAACCCCGCGCGCGCGCGGTGGCGGCC 575
Db |||||
QY 541 GGAGCCTATCGGAGGTGGAGGCGGAGAGTCTTGGGCGGCTGGCAGTGGAGGCGCTGGTG 600
Db |||||
QY 576 GGAGCCTATCGGAGGTGGAGGCGGAGAGTCTTGGGCGGCTGGCAGTGGAGGCGCTGGTG 635
Db |||||
QY 601 AGGACTCGAGGTGCGTGTGCGCGCAGCGCTGGGCGGCAAGCGCGAGGTCTATGTAC 660
Db |||||
QY 636 AGGACTCGAGGTGCGTGTGCGCGCAGCGCTGGGCGGCAAGCGCGAGGTCTATGTAC 695
Db |||||
QY 661 ACAGCCTTCAAGCGCTCGGAGATTTCGTTGAGTACGTGACCTCGGACACA 720
Db |||||
QY 696 ACAGCCTTCAAGCGCTCGGAGATTTCGTTGAGTACGTGACCTCGGACACA 755
Db |||||
QY 721 AGTTGGACCCCATGGGCTGCTGGAGTCTGTCGGGTACTGGAGGAGGACCCCGGGTA 780
Db |||||
QY 756 AGTTGGACCCCATGGGCTGCTGGAGTCTGTCGGGTACTGGAGGAGGACCCCGGGTA 815
Db |||||
QY 781 GGGGCTGTGGTGGGACGTGCGGATCCTTAAACCTCTGGACTCTCGGCTCAGCTTCTTA 840
Db |||||
QY 816 GGGGCTGTGGTGGGATGTGCGGATCCTTAAACCTCTGGACTCTCGGCTCAGCTTCTTA 875
Db |||||
QY 841 AGCAGCCTGCGATATCTGGGTAGCCTTCAATGTGGAGCGGGTGTCTCAGAGCTACTTCCAC 900
Db |||||
QY 876 AGCAGCCTGCGATATCTGGGTAGCCTTCAATGTGGAGCGGGTGTCTCAGAGCTACTTCCAC 935
Db |||||
QY 901 TGTGTATCTGTCATCAGCGGTCTCTAGGCTTATATAGGAATAAATCTTTCAGAGCTTT 960
Db |||||
QY 936 TGTGTATCTGTCATCAGCGGTCTCTAGGCTTATATAGGAATAAATCTTTCAGAGCTTT 995
Db |||||
QY 961 CTTGAGGCTGTGTACAAACAGAGTCTCTGGGTACCCACTGTCATTTTGGGGATGACCGG 1020
Db |||||
QY 996 CTTGAGGCTGTGTACAAACAGAGTCTCTGGGTACCCACTGTCATTTTGGGGATGACCGG 1055
Db |||||
QY 1021 CACCTCACCAACCGCATGCTCAGCATGGTATGCTACCAAGTA 1064
Db |||||
QY 1056 CACCTCACCAACCGCATGCTCAGCATGGTATGCTACCAAGTA 1099
Db |||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 13:46:32 ; Search time 90 Seconds
(without alignments)
1375.148 Million cell updates/sec

Title: US-10-672-399-8

Perfect score: 1697

Sequence: 1 AFLSAHLVAQSLFAYLEHRR.....SGTPGPAATQRRPRPSCGG 320

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980a:*
- 2: Geneseqp1990a:*
- 3: Geneseqp2000a:*
- 4: Geneseqp2001a:*
- 5: Geneseqp2002a:*
- 6: Geneseqp2003a:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1553.5	91.5	578	2	AAW26765 Human hya
2	1546.5	91.1	543	2	AAW36503 Human hya
3	1508.5	88.9	582	3	AAV78135 Recombina
4	1508.5	88.9	583	2	AAW30704 Mouse hya
5	1508.5	88.9	583	2	AAV32503 Hyaluroa
6	1508.5	88.9	583	3	AAV68491 Mouse hya
7	1508.5	88.9	583	3	AAW09948 Murine HA
8	1508.5	88.9	583	6	ABP96028 Mouse hya
9	1508.5	88.9	583	7	AAE39152 Mouse hya
10	1458.5	85.9	563	3	AAV78129 Recombina
11	1362	80.3	584	3	AAV78132 Recombina
12	1358.5	80.1	582	3	AAV78138 Recombina
13	1358.5	80.1	583	3	AAV78131 Recombina
14	1085	63.9	552	3	AAV78133 Recombina
15	1085	63.9	553	3	AAV78136 Recombina
16	965.5	56.9	588	7	ADD93328 Xenopus 1
17	948.5	55.9	552	7	ADC49213 Rabbit hy
18	947	55.8	552	3	AAV78142 Recombina
19	943	55.6	577	3	AAV78128 Recombina
20	942	55.5	552	3	AAV78140 Recombina
21	942	55.5	552	3	AAV78139 Recombina
22	938.5	55.3	553	3	AAV78134 Recombina
23	935.5	55.1	554	2	AAW50010 Murine hy
24	935.5	55.1	554	3	AAV68493 Mouse hya
25	935.5	55.1	554	6	ABP96030 Mouse hya

26	935.5	55.1	554	7	AAE39154 Mouse hya
27	935.5	55.1	554	8	ADJ76214 Marker ge
28	935.5	55.1	557	3	AAV78130 Recombina
29	935	55.1	552	2	AAW50009 Murine hy
30	935	55.1	552	3	AAV68492 Mouse hya
31	935	55.1	552	5	ABW57265 Mouse isc
32	935	55.1	552	6	ABP96029 Mouse hya
33	935	55.1	552	7	ADA49686 Mouse hya
34	935	55.1	552	7	AAE39153 Mouse hya
35	935	55.1	552	7	ADC59299 Mouse HAS
36	935	55.1	553	3	AAV78137 Recombina
37	934	55.0	552	7	ADA49684 Human hya
38	934	55.0	552	7	ADC59297 Human HAS
39	934	55.0	552	7	ADD48818 Human Pro
40	931.5	54.9	573	3	AAV78127 Recombina
41	930	54.8	552	3	AAV78141 Recombina
42	930	54.8	552	7	ADD48816 Rat Prote
43	928	54.7	553	6	ABR48209 Human bla
44	928	54.7	553	6	ABU56490 Lung canc
45	928	54.7	553	6	ABU56686 Lung canc

ALIGNMENTS

RESULT 1
AAW26765
ID AAW26765 standard; protein; 578 AA.
XX
AC AAW26765;
XX
DT 21-MAY-1998 (first entry)
XX
DE Human hyaluronan synthase.
XX
KW Hyaluronan synthase; HAS gene; human; hyaluronic acid; cell adhesion;
KW wound healing; vulnery; tissue repair; scar; keloid; therapy.
XX
OS Homo sapiens.
XX
FH Key
FT Binding-site Location/Qualifiers
FT Modified-site /note= "putative hyaluronan binding site"
FT Modified-site /note= "consensus phosphorylation sequence for protein kinase C"
FT Modified-site /note= "consensus phosphorylation sequence for protein kinase C"
FT Modified-site /note= "consensus phosphorylation sequence for protein kinase C"
FT Modified-site /note= "consensus phosphorylation sequence for cAMP-dependent protein kinase"
XX
WO9740174-A1.
PD 30-OCT-1997.
XX
PF 17-APR-1997; 97WO-US006350.
XX
PR 22-APR-1996; 96US-00635552.
XX
PA (LEUK-) LEUKOSITE INC.
XX
PI Brieskin MJ;
XX
DR WPI; 1997-549359/50.
XX
PT N-PSDB; AAT99541.
XX
PT Human hyaluronan synthase - useful for recombinant production of
PT hyaluronic acid for wound healing, tissue repair and reducing


```

PD 11-JAN-2000.
XX
XX
XX 24-JUN-1998; 98JP-00193788.
XX
XX 24-JUN-1998; 98JP-00193788.
XX
XX (SEK ) SEIKAGAKU KOGYO CO LTD.
XX
XX WPI; 2000-140125/13.
XX
XX A hyaluronate synthase modified protein - useful as a research reagent
XX for biochemical research and medical development.
XX
XX Claim 10; Page; 30pp; Japanese.
XX
XX The present invention describes a recombinant protein which consists of
XX three continuous regions (N-terminal region, internal region and C-
XX terminal region) where one or two regions among the above three regions
XX is selected from the three hyaluronate synthase (HAS) modified proteins
XX HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
XX HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
XX is useful as a research reagent for biochemical research and medical
XX development. The invention provides a HAS modified protein of increased
XX or lowered activity. AAY78127 to AAY78142 represent specifically claimed
XX examples of recombinant proteins from the present invention. N.B. The
XX present sequence is not given in the present specification, but is
XX derived from sequences given as specified in the claim
XX
XX Sequence 582 AA;
XX
XX Query Match 88.9%; Score 1508.5; DB 3; Length 582;
XX Best Local Similarity 89.9%; Pred. No. 3.5e-167;
XX Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;
XX
XX 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
XX 60 AFLSAHLVAQSLFAYLEHRRVAAAARRSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
XX
XX 56 ASARALLYPRARLRVLMVVDGNRAEDLYMVDMFREVFADEDPATYVMDGNVHQWPEPA-A 114
XX 120 TSARALLYPHTRLRVLMVVDGNRAEDLYMVDMFREVFADEDPATYVMDGNVHQWPEPAE 179
XX
XX 115 AGAVGAGAYREVEADPGRLAVALVTRTRCVCVAQRWGKREVMYTAFAKALGDSVDYVQ 174
XX 180 TGAEGEGAYREVEADPGRLAVALVTRTRCVCVAQRWGKREVMYTAFAKALGDSVDYVQ 239
XX
XX 175 VCDSDTRLDPMALLEVRVLDDEPRVGVAGGVDVRLNPLDSWVSLSLRYVWAFNVERA 234
XX 240 VCDSDTRLDPMALLEVRVLDDEPRVGVAGGVDVRLNPLDSWVSLSLRYVWAFNVERA 299
XX
XX 235 CQSYFHCVCISGPGGLYRNLLQOFLAAYNQKFLGTHCTFGDDRHLTNRLMSGYATK 294
XX 300 CQSYFHCVCISGPGGLYRNLLQOFLAAYNQKFLGTHCTFGDDRHLTNRLMSGYATK 359
XX
XX 295 -----AEGTRWS 301
XX 360 YTSRRCYSETPSSFLRWLSQOTRWS 385
XX
XX RESULT 4
XX AAW30704
XX ID AAW30704 standard; protein; 583 AA.
XX
XX AC AAW30704;
XX
XX XX 17-OCT-2003 (revised)
XX DT 19-DEC-1997 (first entry)
XX
XX DE Mouse hyaluronate synthase enzyme.
XX
XX XX Hyaluronate synthase; hyaluronic acid; polysaccharide.
XX
XX OS Mus sp; (strain FM3A).
XX
XX
XX JP09224674-A.
XX
XX 02-SEP-1997.
XX
XX 26-FEB-1996; 96JP-00038336.
XX
XX 26-FEB-1996; 96JP-00038336.
XX
XX (KAGG ) KAGAKU GIJUTSUCHO CHOKAN KANBO.
XX
XX WPI; 1997-484102/45.
XX
XX N-PADB; AAT91655.
XX
XX Hyaluronate synthase isolated from mouse cells - useful for large-scale
XX production of hyaluronic acid.
XX
XX Claim 4; Page 13-14; 15pp; Japanese.
XX
XX A mouse-derived cell strain, FM3A P15-A, was cultured in Eagle minimal
XX medium containing 10 % heat-inactivated bovine serum, twice concentration
XX of amino acids and vitamins and penicillin and streptomycin at 37 degrees
XX Celsius. The culture was subjected to immobilised erythrocyte exclusion
XX to examine the extent of extracellular formation of hyaluronic acid
XX matrix. Cells which showed high formation were recovered and named FM3A
XX HAI. Cells synthesising no hyaluronic acid were recovered and named HAS-.
XX The HAS- cell in which polyoma large T antigen was expressed was prepared
XX (HAS- P cell ). Poly(A)+ RNA was isolated from FM3A HAI and cDNA was
XX prepared and was used for constructing a library in HAS- P cells. Cells
XX having hyaluronic acid synthetic activity were selected from the
XX transformants and plasmid DNA was recovered and amplified in E.coli. The
XX resulting genomic DNA sequence codes for hyaluronate synthase having the
XX present sequence, which is used for large-scale production of hyaluronic
XX acid, a component of drugs and cosmetics. (Updated on 17-OCT-2003 to
XX standardise OS field)
XX
XX Sequence 583 AA;
XX
XX Query Match 88.9%; Score 1508.5; DB 2; Length 583;
XX Best Local Similarity 89.9%; Pred. No. 3.6e-167;
XX Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;
XX
XX 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
XX 60 AFLSAHLVAQSLFAYLEHRRVAAAARRSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
XX
XX 56 ASARALLYPRARLRVLMVVDGNRAEDLYMVDMFREVFADEDPATYVMDGNVHQWPEPA-A 114
XX 120 TSARALLYPHTRLRVLMVVDGNRAEDLYMVDMFREVFADEDPATYVMDGNVHQWPEPAE 179
XX
XX 115 AGAVGAGAYREVEADPGRLAVALVTRTRCVCVAQRWGKREVMYTAFAKALGDSVDYVQ 174
XX 180 TGAEGEGAYREVEADPGRLAVALVTRTRCVCVAQRWGKREVMYTAFAKALGDSVDYVQ 239
XX
XX 175 VCDSDTRLDPMALLEVRVLDDEPRVGVAGGVDVRLNPLDSWVSLSLRYVWAFNVERA 234
XX 240 VCDSDTRLDPMALLEVRVLDDEPRVGVAGGVDVRLNPLDSWVSLSLRYVWAFNVERA 299
XX
XX 235 CQSYFHCVCISGPGGLYRNLLQOFLAAYNQKFLGTHCTFGDDRHLTNRLMSGYATK 294
XX 300 CQSYFHCVCISGPGGLYRNLLQOFLAAYNQKFLGTHCTFGDDRHLTNRLMSGYATK 359
XX
XX 295 -----AEGTRWS 301
XX 360 YTSRRCYSETPSSFLRWLSQOTRWS 385
XX
XX RESULT 5
XX AAY32503
XX ID AAY32503 standard; protein; 583 AA.
XX
XX AC AAY32503;
XX
XX XX

```


DT 19-OCT-1999 (first entry)
 XX Hyaluronate synthase protein sequence.
 DE Hyaluronate synthase; promoter; mouse; cell growth inhibitor.
 KW Mus sp.
 OS JP11196875-A.
 XX 27-JUL-1999.
 XX 14-JAN-1998; 98JP-00006191.
 XX 14-JAN-1998; 98JP-00006191.
 XX (SEKG) SEIKAGAKU KOGYO CO LTD.
 XX WPI; 1999-496653/42.
 DR N-PSDB; AAZ10862.
 XX New promoter DNA of hyaluronate synthase - used to specifically express
 PT gene with cell growth inhibiting activity.
 XX Example 1; Page 8-10; 13pp; Japanese.
 XX This sequence is encoded by the mouse hyaluronate synthase gene, which
 CC was used to isolate the hyaluronate synthase promoter of the invention.
 CC The promoter can be used for specifically expressing a gene having cell
 CC growth inhibiting activity
 XX Sequence 583 AA;

KW research reagent; biochemical research; medical development.
 XX Mus sp.
 OS JP2000004886-A.
 XX 11-JAN-2000.
 XX 24-JUN-1998; 98JP-00193788.
 XX 24-JUN-1998; 98JP-00193788.
 XX (SEKG) SEIKAGAKU KOGYO CO LTD.
 XX WPI; 2000-140125/13.
 DR N-PSDB; AAZ88199.
 XX A hyaluronate synthase modified protein - useful as a research reagent
 PT for biochemical research and medical development.
 XX Claim 1; Page 14-16; 30pp; Japanese.
 XX The present invention describes a recombinant protein which consists of
 CC three continuous regions (N-terminal region, internal region and C-
 CC terminal region) where one or two regions among the above three regions
 CC is selected from the three hyaluronate synthase (HAS) modified proteins
 CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
 CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
 CC is useful as a research reagent for biochemical research and medical
 CC development. The invention provides a HAS modified protein of increased
 CC or lowered activity. The present sequence represents mouse HAS1
 XX Sequence 583 AA;

Query Match 88.9%; Score 1508.5; DB 2; Length 583;
 Best Local Similarity 89.9%; Pred. No. 3.6e-167;
 Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;
 QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
 DB 60 AFLSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
 QY 56 ASARALLYPRARLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPA-A 114
 DB 120 TSARALLYPHTRLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPAE 179
 QY 115 AGAVGAGAYREVEAEDPGRLAVALVTRRCVCVCAQRWGGKREVMYTAFAKALGDSVDYVQ 174
 DB 180 TGAVGEGAYREVEAEDPGRLAVALVTRRCVCVCAQRWGGKREVMYTAFAKALGDSVDYVQ 239
 QY 175 VCDSDTRLDPMALLELVRVLDDEDPRVGAVGGDVRLINPLDSWVSLSLRYWVAFNVERA 234
 DB 240 VCDSDTRLDPMALLELVRVLDDEDPRVGAVGGDVRLINPLDSWVSLSLRYWVAFNVERA 299
 QY 235 CQSYFHCVCISGGLYRNLLQOFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQ 294
 DB 300 CQSYFHCVCISGGLYRNLLQOFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQ 359
 QY 295 -----AEGTRWS 301
 DB 360 YTSRRCYSETPSSFLRWLSQOTRWS 385

RESULT 6
 AAY68491
 ID AAY68491 standard; protein; 583 AA.
 XX AC AAY68491;
 XX 27-APR-2000 (first entry)
 DT Mouse hyaluronate synthase modified protein HAS1.
 DE Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
 KW

RESULT 7
 AAB09948
 ID AAB09948 standard; protein; 583 AA.
 XX AC AAB09948;
 XX 19-OCT-2000 (first entry)
 DT Murine HAS1 protein.
 DE
 XX

Query Match 88.9%; Score 1508.5; DB 3; Length 583;
 Best Local Similarity 89.9%; Pred. No. 3.6e-167;
 Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;
 QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
 DB 60 AFLSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
 QY 56 ASARALLYPRARLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPA-A 114
 DB 120 TSARALLYPHTRLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPAE 179
 QY 115 AGAVGAGAYREVEAEDPGRLAVALVTRRCVCVCAQRWGGKREVMYTAFAKALGDSVDYVQ 174
 DB 180 TGAVGEGAYREVEAEDPGRLAVALVTRRCVCVCAQRWGGKREVMYTAFAKALGDSVDYVQ 239
 QY 175 VCDSDTRLDPMALLELVRVLDDEDPRVGAVGGDVRLINPLDSWVSLSLRYWVAFNVERA 234
 DB 240 VCDSDTRLDPMALLELVRVLDDEDPRVGAVGGDVRLINPLDSWVSLSLRYWVAFNVERA 299
 QY 235 CQSYFHCVCISGGLYRNLLQOFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQ 294
 DB 300 CQSYFHCVCISGGLYRNLLQOFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQ 359
 QY 295 -----AEGTRWS 301
 DB 360 YTSRRCYSETPSSFLRWLSQOTRWS 385

Query Match 88.9%; Score 1508.5; DB 3; Length 583;
 Best Local Similarity 89.9%; Pred. No. 3.6e-167;
 Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;
 QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
 DB 60 AFLSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
 QY 56 ASARALLYPRARLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPA-A 114
 DB 120 TSARALLYPHTRLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPAE 179
 QY 115 AGAVGAGAYREVEAEDPGRLAVALVTRRCVCVCAQRWGGKREVMYTAFAKALGDSVDYVQ 174
 DB 180 TGAVGEGAYREVEAEDPGRLAVALVTRRCVCVCAQRWGGKREVMYTAFAKALGDSVDYVQ 239
 QY 175 VCDSDTRLDPMALLELVRVLDDEDPRVGAVGGDVRLINPLDSWVSLSLRYWVAFNVERA 234
 DB 240 VCDSDTRLDPMALLELVRVLDDEDPRVGAVGGDVRLINPLDSWVSLSLRYWVAFNVERA 299
 QY 235 CQSYFHCVCISGGLYRNLLQOFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQ 294
 DB 300 CQSYFHCVCISGGLYRNLLQOFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQFLQ 359
 QY 295 -----AEGTRWS 301
 DB 360 YTSRRCYSETPSSFLRWLSQOTRWS 385

KW Murine; HAS1; hyaluronate synthase; marker gene; intracellular loop;
 KW Gene targeting; ss.
 XX Mus sp.
 XX JP2000116382-A.
 XX 25-APR-2000.
 XX 13-OCT-1998; 98JP-00291201.
 XX 13-OCT-1998; 98JP-00291201.
 XX (SEK) SEIKAGAKU KOGYO CO LTD.
 XX WPI; 2000-369404/32.
 XX N-PSDB; AAA39987.
 XX DNA used for gene targeting of the hyaluronate synthase gene (HAS1).
 XX Example; Page 10-11; 14pp; Japanese.
 XX This invention describes a novel DNA which contains an heterologous DNA
 CC (I) to be introduced to chromosomal DNA of a host cell, a first and
 CC second homologous region DNA (II) connected respectively to the 5' and 3'
 CC sides of the introduced DNA, and a negative marker gene (III) expressable
 CC in the host cell. (I) contains a positive marker gene expressible in the
 CC host cell. (II) and the region encoding the intracellular loop of
 CC the hyaluronate synthase (HAS) 1 protein from chromosomal DNA is
 CC constituted so as to cause a homologous recombination. The DNA can be
 CC used for gene targeting of the HAS1 gene. This sequence represents the
 CC murine HAS1 protein described in the method of the invention
 XX Sequence 583 AA;
 XX
 Query Match 88.9%; Score 1508.5; DB 3; Length 583;
 Best Local Similarity 89.9%; Pred. No. 3.6e-167;
 Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;
 QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
 Db 60 AFLSAHLVAQSLFAYLEHRRVAAAARRSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
 QY 56 ASARALLYPRARLRVLMVVDGNRAEDLYMDMFREVFADEDPATYVMDGNTHQHPWEPA-A 114
 Db 120 TSARALLYPHTRLRVLMVVDGNRAEDLYMDMFREVFADEDPATYVMDGNTHQHPWEPAE 179
 QY 115 AGAUCAGAYREVEAEDPGRLAVALVTRRCVCVQAQRWGGKREVMYTFKALGDSVDVYVQ 174
 Db 180 TGAVGEGAYREVEAEDPGRLAVALVTRRCVCVQAQRWGGKREVMYTFKALGDSVDVYVQ 239
 QY 175 VCDSDTRLDPMALLELRVLDSDPRVGAAGVDVRLNPLDSWVSFLSSLYVWAFNVERA 234
 Db 240 VCDSDTRLDPMALLELRVLDSDPRVGAAGVDVRLNPLDSWVSFLSSLYVWAFNVERA 299
 QY 235 CQSYFHCVCISGSLGLYRNLLQQLFLEAWYNQKFLGTHCTFGDDRHLTNRLMSGYATK 294
 Db 300 CQSYFHCVCISGSLGLYRNLLQQLFLEAWYNQKFLGTHCTFGDDRHLTNRLMSGYATK 359
 QY 295 -----AEGTRWS 301
 Db 360 YTSRRCYSETPSSFLRLMSQOTRWS 385
 RESULT 8
 ID ABP96028
 XX standard; protein; 583 AA.
 AC ABP96028;
 XX
 DT 01-APR-2000 (first entry)
 XX
 DE Mouse hyaluronan synthase 1 protein SEQ ID NO:4.

XX Mouse; hyaluronan synthase; HAS; HAS1; enzyme; ophthalmological;
 KW osteopathic; antiarthritic; gene therapy; angiogenesis inhibitor;
 KW hyaluronic acid; dihydrazide; bioconjugate; dry eye syndrome; eye;
 KW osteoarthritis.
 XX Mus sp.
 XX WO2003006068-A1.
 XX 23-JAN-2003.
 XX 10-JUL-2001; 2001WO-US021785.
 XX 10-JUL-2001; 2001WO-US021785.
 XX (CLEA-) CLEAR SOLUTIONS BIOTECH INC.
 XX Dehazya P, Chen W;
 XX WPI; 2003-221664/21.
 XX N-PSDB; ABZ76734.
 XX Novel dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate
 PT for treating dry eye syndrome, has derivatized hyaluronic acid
 PT crosslinked to nucleic acid encoding protein with hyaluronan synthase
 PT activity.
 XX Claim 20; Page 58-59; 62pp; English.
 XX The present invention describes a dihydrazide derivatized hyaluronic acid
 CC (HA)/nucleic acid (NA) bioconjugate (I), with derivatized HA linked to NA
 CC encoding protein with hyaluronan synthase (HAS) activity, where NA has
 CC sequence with 79-85% sequence identity (SI) to a sequence of 1752, 1659
 CC or 1665 nucleotides (see ABZ76734 to ABZ76736) or a sequence encoding
 CC proteins with 66-90% SI to a sequence of 583, 552 or 554 amino acids (see
 CC ABP96028 to ABP96030). (I) has ophthalmological, osteopathic and
 CC antiarthritic activities, and can be used in gene therapy and as an
 CC inhibitor of angiogenesis, and as an inducer of expression of (HA) in
 CC human corneal epithelial cell. (I) is useful for transfecting a cell of
 CC an eye with nucleic acid, where the nucleic acid comprises (ABZ76734 to
 CC ABZ76736), by contacting the cell with (I) comprising the nucleic acid.
 CC (I) is useful for treating dry eye syndrome in an individual. (I) is
 CC useful in gene therapy applications for the treatment of a variety of
 CC medical conditions including dry eye syndrome or other medical conditions
 CC where an increase in the production of (HA) in the eye would be
 CC therapeutic (e.g., osteoarthritis of the articular joints). (I) is also
 CC useful for inhibiting angiogenesis for the treatment of macular
 CC degeneration or genes related to lipid biosynthesis that helps to restore
 CC the lipid component of the tear film, and as reagents for in vitro
 CC transformation of any cell, preferably a eukaryotic cell, more preferably
 CC a human eye cell. The present sequence represents mouse HAS1 which is
 CC used in the exemplification of the present invention
 XX Sequence 583 AA;
 XX
 Query Match 88.9%; Score 1508.5; DB 6; Length 583;
 Best Local Similarity 89.9%; Pred. No. 3.6e-167;
 Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;
 QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
 Db 60 AFLSAHLVAQSLFAYLEHRRVAAAARRSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
 QY 56 ASARALLYPRARLRVLMVVDGNRAEDLYMDMFREVFADEDPATYVMDGNTHQHPWEPA-A 114
 Db 120 TSARALLYPHTRLRVLMVVDGNRAEDLYMDMFREVFADEDPATYVMDGNTHQHPWEPAE 179
 QY 115 AGAUCAGAYREVEAEDPGRLAVALVTRRCVCVQAQRWGGKREVMYTFKALGDSVDVYVQ 174
 Db 180 TGAVGEGAYREVEAEDPGRLAVALVTRRCVCVQAQRWGGKREVMYTFKALGDSVDVYVQ 239
 QY 175 VCDSDTRLDPMALLELRVLDSDPRVGAAGVDVRLNPLDSWVSFLSSLYVWAFNVERA 234

```
Db 240 VCDSDTRLDPALLELVRLVDEDPRVGAVGGDVRLINPLDSWVSFLSLRTWVAFNVERA 299
QY 235 CQSYFHCVCISGPIGLYRNLLQOFLBAWYNQKFLGTHCTFGDDRHLTNRMLSMGYATK 294
Db 300 CQSYFHCVCISGPIGLYRNLLQOFLBAWYNQKFLGTHCTFGDDRHLTNRMLSMGYATK 359
QY 295 -----AEGTRWS 301
Db 360 YTSRRCYSETPSSFLRWLSQOTRWS 385

RESULT 9
AAE39152
ID AAE39152 standard; protein; 583 AA.
XX
AC AAE39152;
XX
DT 18-DEC-2003 (first entry)
XX
DE Mouse hyaluronan synthase (HAS) 1.
XX
KW Mouse; hyaluronic acid; HA; hyaluronan synthase; HAS; dry eye syndrome;
KW eye cell; osteoarthritis; gene therapy; enzyme.
XX
OS Mus sp.
XX
FN US2003087850-A1.
XX
PD 08-MAY-2003.
XX
PF 10-JUL-2001; 2001US-00902939.
XX
PR 10-JUL-2001; 2001US-00902939.
XX
PA (DEHA/) DEHAZYA P.
PA (CHEN/) CHEN W.
XX
PI Debazyia P, Chen W;
XX
DR N-PSDB; AAD59442.
XX
PT Dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate
PT comprising derivatized hyaluronic acid cross linked to nucleic acid
PT encoding hyaluronan synthase useful for treating dry eye syndrome.
XX
PS Claim 18; Page 15-16; 31pp; English.
XX
CC The invention relates to dihydrazide derivatised hyaluronic acid (HA)/
CC nucleic acid bioconjugate comprising derivatised HA cross linked to
CC nucleic acid encoding hyaluronan synthase (HAS). The invention is useful
CC for transfected an eye cell of an individual. It is useful for treating
CC dry eye syndrome and osteoarthritis of the particular joints. The
CC invention is also useful in gene therapy. The present sequence is mouse
CC HAS1 enzyme
XX
SQ Sequence 583 AA;
Query Match 88.9%; Score 1508.5; DB 7; Length 583;
Best Local Similarity 89.9%; Pred. No. 3.6e-167; Indels 25; Gaps 3;
Matches 293; Conservative 2; Mismatches 6;
QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
Db 60 AFLSAHLVAQSLFAYLEHRRVAAAARRSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
QY 56 ASARALLYPRALRLVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNVHQPWEPA-A 114
Db 120 TSARALLYPHTRLRLVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNVHQPWEPAEA 179
QY 115 AGAVGAGAYREAEADPGRLAVEALVRTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQ 174
|||||
```

```
Db 180 TGAVGEGAYREAEADPGRLAVEALVRTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQ 239
QY 175 VCDSDTRLDPALLELVRLVDEDPRVGAVGGDVRLINPLDSWVSFLSLRTWVAFNVERA 234
Db 240 VCDSDTRLDPALLELVRLVDEDPRVGAVGGDVRLINPLDSWVSFLSLRTWVAFNVERA 299
QY 235 CQSYFHCVCISGPIGLYRNLLQOFLBAWYNQKFLGTHCTFGDDRHLTNRMLSMGYATK 294
Db 300 CQSYFHCVCISGPIGLYRNLLQOFLBAWYNQKFLGTHCTFGDDRHLTNRMLSMGYATK 359
QY 295 -----AEGTRWS 301
Db 360 YTSRRCYSETPSSFLRWLSQOTRWS 385

RESULT 10
AAAY78129
ID AAY78129 standard; protein; 563 AA.
XX
AC AAY78129;
XX
DT 27-APR-2000 (first entry)
XX
DE Recombinant chimeric hyaluronate synthase modified protein #3.
XX
KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX
OS Mus sp.
OS Synthetic.
OS Chimeric.
XX
FN JP2000004886-A.
XX
PD 11-JAN-2000.
XX
PE 24-JUN-1998; 98JP-00193788.
XX
PR 24-JUN-1998; 98JP-00193788.
XX
PA (SEKG ) SEIKAGAKU KOGYO CO LTD.
XX
DR WPI; 2000-140125/13.
XX
FT A hyaluronate synthase modified protein - useful as a research reagent
FT for biochemical research and medical development.
XX
PS Claim 4; Page; 30pp; Japanese.
XX
CC The present invention describes a recombinant protein which consists of
CC three continuous regions (N-terminal region, internal region and C-
CC terminal region) where one or two regions among the above three regions
CC is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. AAY78127 to AAY78142 represent specifically claimed
CC examples of recombinant proteins from the present invention. N.B. The
CC present sequence is not given in the present specification, but is
CC derived from sequences given as specified in the claim
XX
SQ Sequence 563 AA;
Query Match 85.9%; Score 1458.5; DB 3; Length 563;
Best Local Similarity 89.2%; Pred. No. 2.5e-161;
Matches 282; Conservative 3; Mismatches 6; Indels 25; Gaps 3;
QY 11 SLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPR 65
Db 51 ALFAYLEHRRVAAAARRSLAKGPLDAATARSVALTISAYQEDPAYLRQCLTSARALLYPH 110
QY 66 ARLRLVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNVHQPWEPA-AAGAVGAGAYR 124
|||||
```

Db 111 TRLRLVMVVDGNRAEDLYVMDMFRVFADEDPATYVMDGNYPWEPAEATGAVGEGAYR 170
QY 125 EVEADPGRLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDTRLDLP 184
Db 171 EVEADPGRLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDTRLDLP 230
QY 185 MALLELRVLDDEPRVGAAGVGDVRIINPLDSWVSFLSSLRVYVAFNVERACOSYFHCYSC 244
Db 231 MALLELRVLDDEPRVGAAGVGDVRIINPLDSWVSFLSSLRVYVAFNVERACOSYFHCYSC 290
QY 245 TSGPLGLYRNLLQOFLFLEAWYVQKFLGTHCTFGDDRHLTNRLMSMGYATK----- 294
Db 291 TSGPLGLYRNLLQOFLFLEAWYVQKFLGTHCTFGDDRHLTNRLMSMGYATKYTSRRCYSE 350
QY 295 -----AEGTRWS 301
Db 351 TPSSFRLMSQOTRWS 366

RESULT 11
AAV78132
ID AAV78132 standard; protein; 584 AA.
XX
AC AAV78132;
XX
DT 27-APR-2000 (first entry)
XX
DE Recombinant chimeric hyaluronate synthase modified protein #6.
XX
KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX
OS Mus sp.
OS Synthetic.
OS Chimeric.
XX JP2000004886-A.
XX
PD 11-JAN-2000.
XX
PF 24-JUN-1998; 98JP-00193788.
XX
PR 24-JUN-1998; 98JP-00193788.
XX
PA (SEKG) SEIKAGAKU KOGYO CO LTD.
XX
DR WPI; 2000-140125/13.
XX
PT A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.
XX
PS Claim 7; Page; 30pp; Japanese.
XX

CC The present invention describes a recombinant protein which consists of
CC three continuous regions (N-terminal region, internal region and C-
CC terminal region) where one or two regions among the above three regions
CC is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. AAV78127 to AAV78142 represent specifically claimed
CC examples of recombinant proteins from the present invention. N.B. The
CC present sequence is not given in the present specification, but is
CC derived from sequences given as specified in the claim

XX Sequence 584 AA;

Query Match 80.3%; Score 1362; DB 3; Length 584;
Best Local Similarity 80.4%; Pred. No. 5.6e-150;
Matches 263; Conservative 19; Mismatches 19; Indels 26; Gaps 4;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
Db 60 AFLSAHLVAQSLFAYLEHRRVAAAARSLAKGFLDAATARSVALTISAYQEDPAYLRQCL 119
QY 56 ASARALLYPRLRLRLVLMVVDGNRAEDLYVMDMFRVFADEDPATYVMDGNYPWEPAE-A 114
Db 120 TSARALLYPHTRLRLVLMVVDGNRAEDLYVMDMFRVFADEDPATYVMDGNYPWEPAE-A 179
QY 115 AGAVGAGAYREVEADPGRLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYV 173
Db 180 TGAVGEGAYREVEADPGRLAVALVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYV 239
QY 174 QVCDSDTRLDPMALLELRVLDDEPRVGAAGVGDVRIINPLDSWVSFLSSLRVYVAFNVER 233
Db 240 QVCDSDTVLDPACTIEMLRVLEEDPQVGVGSDVQILNKYDSWISFLSSVRVYMAFNVER 299
QY 234 ACOSYFHCYSCISGPLGLYRNLLQOFLFLEAWYVQKFLGTHCTFGDDRHLTNRLMSMGYAT 293
Db 300 ACOSYFHCYSCISGPLGLYRNLLQOFLFLEAWYVQKFLGTHCTFGDDRHLTNRLMSMGYAT 359
QY 294 K-----AEGTRWS 301
Db 360 KYTARSKCLTETPTTRYLRWLNQOTRWS 386

RESULT 12
AAV78138
ID AAV78138 standard; protein; 582 AA.
XX
AC AAV78138;
XX
DT 27-APR-2000 (first entry)
XX
DE Recombinant chimeric hyaluronate synthase modified protein #12.
XX
KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX
OS Mus sp.
OS Synthetic.
OS Chimeric.
XX JP2000004886-A.
XX
PD 11-JAN-2000.
XX
PF 24-JUN-1998; 98JP-00193788.
XX
PR 24-JUN-1998; 98JP-00193788.
XX
PA (SEKG) SEIKAGAKU KOGYO CO LTD.
XX
DR WPI; 2000-140125/13.
XX
PT A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.
XX
PS Claim 13; Page; 30pp; Japanese.

CC The present invention describes a recombinant protein which consists of
CC three continuous regions (N-terminal region, internal region and C-
CC terminal region) where one or two regions among the above three regions
CC is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. AAV78127 to AAV78142 represent specifically claimed
CC examples of recombinant proteins from the present invention. N.B. The
CC present sequence is not given in the present specification, but is
CC derived from sequences given as specified in the claim

XX Sequence 582 AA;

Query Match 80.1%; Score 1358.5; DB 3; Length 582;
Best Local Similarity 79.4%; Pred. No. 1.4e-149;
Matches 259; Conservative 21; Mismatches 21; Indels 25; Gaps 3;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
DB 60 AFLSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
QY 56 ASARALLYPRARLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQPWEPA-A 114
DB 120 TSARALLYPHTRLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQPWEPAEA 179
QY 115 AGAVGAGAYREVEAEDPGRLAVEALVRTRRCVCAQWRGKREVMYTAFAKALGDSVDVYQ 174
DB 180 TGAVGEGAYREVEAEDPGRLAVEALVRTRRCVCAQWRGKREVMYTAFAKALGDSVDVYQ 239
QY 175 VCDSDTRLDPMALLVRVLDPRVGAAGVGDVRLNPLDSWVSFLSLRYVWAFNVERA 234
DB 240 VCDSDTRLDPMALLVRVLDPRVGAAGVGDVRLNPLDSWVSFLSLRYVWAFNVERA 299
QY 235 COSYFHCVCISGPIGLYRNLLQOFLBAWYNQKFLGTHCTFGDDRHLLTNRLMWSGYATK 294
DB 300 COSYFHCVCISGPIGLYRNLLQOFLBAWYNQKFLGTHCTFGDDRHLLTNRLMWSGYATK 359
QY 295 -----AEGTRWS 301
DB 360 YTARSKCLTETPIEYLRWLNQOTRWS 385

RESULT 13
AAV78131
ID AAV78131 standard; protein; 583 AA.
AC AAV78131;
XX
XX
XX
XX 27-APR-2000 (first entry)
DE Recombinant chimeric hyaluronate synthase modified protein #5.
XX
XX
XX Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX
XX
XX Mus sp.
OS Synthetic.
OS Chimeric.
XX
XX JP2000004886-A.
XX
XX PD 11-JAN-2000.
XX
XX PF 24-JUN-1998; 98JP-00193788.
XX
XX PR 24-JUN-1998; 98JP-00193788.
XX
XX PA (SEK) SEIKAGAKU KOGYO CO LTD.
XX
XX DR WPI; 2000-140125/13.
XX
XX A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.
XX
XX PS Claim 6; Page; 30pp; Japanese.

The present invention describes a recombinant protein which consists of three continuous regions (N-terminal region, internal region and C-terminal region) where one or two regions among the above three regions is selected from the three hyaluronate synthase (HAS) modified proteins HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71, HAS2 amino acids 52 to 379 and HAS3 amino acids 410 to 583). The protein is useful as a research reagent for biochemical research and medical development. The invention provides a HAS modified protein of increased or lowered activity. AAV78127 to AAV78142 represent specifically claimed

CC examples of recombinant proteins from the present invention. N.B. The
CC present sequence is not given in the present specification, but is
CC derived from sequences given as specified in the claim

XX SQ Sequence 583 AA;

Query Match 80.1%; Score 1358.5; DB 3; Length 583;
Best Local Similarity 79.4%; Pred. No. 1.4e-149;
Matches 259; Conservative 21; Mismatches 21; Indels 25; Gaps 3;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
DB 60 AFLSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
QY 56 ASARALLYPRARLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQPWEPA-A 114
DB 120 TSARALLYPHTRLRVLMVVDGNRAEDLYWDMFREVFADEDPATYVWDGNYHQPWEPAEA 179
QY 115 AGAVGAGAYREVEAEDPGRLAVEALVRTRRCVCAQWRGKREVMYTAFAKALGDSVDVYQ 174
DB 180 TGAVGEGAYREVEAEDPGRLAVEALVRTRRCVCAQWRGKREVMYTAFAKALGDSVDVYQ 239
QY 175 VCDSDTRLDPMALLVRVLDPRVGAAGVGDVRLNPLDSWVSFLSLRYVWAFNVERA 234
DB 240 VCDSDTRLDPMALLVRVLDPRVGAAGVGDVRLNPLDSWVSFLSLRYVWAFNVERA 299
QY 235 COSYFHCVCISGPIGLYRNLLQOFLBAWYNQKFLGTHCTFGDDRHLLTNRLMWSGYATK 294
DB 300 COSYFHCVCISGPIGLYRNLLQOFLBAWYNQKFLGTHCTFGDDRHLLTNRLMWSGYATK 359
QY 295 -----AEGTRWS 301
DB 360 YTARSKCLTETPIEYLRWLNQOTRWS 385

RESULT 14
AAV78133
ID AAV78133 standard; protein; 552 AA.
AC AAV78133;
XX
XX
XX 27-APR-2000 (first entry)
DE Recombinant chimeric hyaluronate synthase modified protein #7.
XX
XX Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
KW research reagent; biochemical research; medical development; chimeric.
XX
XX Mus sp.
OS Synthetic.
OS Chimeric.
XX
XX JP2000004886-A.
XX
XX PD 11-JAN-2000.
XX
XX PF 24-JUN-1998; 98JP-00193788.
XX
XX PR 24-JUN-1998; 98JP-00193788.
XX
XX PA (SEK) SEIKAGAKU KOGYO CO LTD.
XX
XX DR WPI; 2000-140125/13.

A hyaluronate synthase modified protein - useful as a research reagent for biochemical research and medical development.

Claim 8; Page; 30pp; Japanese.

The present invention describes a recombinant protein which consists of three continuous regions (N-terminal region, internal region and C-terminal region) where one or two regions among the above three regions is selected from the three hyaluronate synthase (HAS) modified proteins

CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. AAV78127 to AAV78142 represent specifically claimed
CC examples of recombinant proteins from the present invention. N.B. The
CC present sequence is not given in the present specification, but is
CC derived from sequences given as specified in the claim

XX Sequence 552 AA;

Query Match 63.9%; Score 1085; DB 3; Length 552;
Best Local Similarity 65.6%; Pred. No. 1.7e-117;
Matches 210; Conservative 31; Mismatches 45; Indels 34; Gaps 5;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
DB 51 AFLASHLLIQSLFAFLEHRRKMKKSLETPI--KLNTVALCIAAYQEDPDYLRKCLQSVKR 108
QY 61 LLYPRARLVLMVVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQFWPEPAAAGAVGA 120
DB 109 LLYPRARLVLMVVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQFWPEPAAAGAVGA 120
QY 121 GAYREVEAEDPGLAVEALVTRRCVCVAQRWGGKREVMYTAFAKALGSDVDYVQVCDSDT 180
DB 158 GETEESHKESQH--VTQLVLSNKSICIMQKGGKREVMYTAFAKALGSDVDYVQVCDSDT 215
QY 181 RLDPMALLELVRLVLEDDPRVAGVGDVRLNPLDSWVSFLSLRYWVAFNVERACQSYFH 240
DB 216 RLDPMALLELVRLVLEDDPRVAGVGDVRLNPLDSWVSFLSLRYWVAFNVERACQSYFH 275
QY 241 CVSCISGPLGLYRNLLQOQFLEAWYKFLGTHCTFGDDRHLTNRMLSMGYATK----- 294
DB 276 CVSCISGPLGLYRNLLQOQFLEAWYKFLGTHCTFGDDRHLTNRMLSMGYATK----- 335
QY 295 -----AEGTRWS 301
DB 336 CYSETPSSFLRWLSQQTWS 355

RESULT 15

ID AAY78136 standard; protein; 553 AA.

XX AC AAY78136;

XX DT 27-APR-2000 (first entry)

XX DE Recombinant chimeric hyaluronate synthase modified protein #10.

XX KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
XX research reagent; biochemical research; medical development; chimeric.

OS Mus sp.
OS Synthetic.
OS Chimeric.

XX JN JP2000004886-A.

XX PD 11-JAN-2000.

XX PP 24-JUN-1998; 98JP-00193788.

XX PR 24-JUN-1998; 98JP-00193788.

XX PA (SEGK) SEIKAGAKU KOGYO CO LTD.

XX DR WPI; 2000-140125/13.

XX PT A hyaluronate synthase modified protein - useful as a research reagent
XX for biochemical research and medical development.

XX PS Claim 11; Page; 30pp; Japanese.

XX

CC The present invention describes a recombinant protein which consists of
CC three continuous regions (N-terminal region, internal region and C-
CC terminal region) where one or two regions among the above three regions
CC is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. AAV78127 to AAV78142 represent specifically claimed
CC examples of recombinant proteins from the present invention. N.B. The
CC present sequence is not given in the present specification, but is
CC derived from sequences given as specified in the claim

XX Sequence 553 AA;

Query Match 63.9%; Score 1085; DB 3; Length 553;
Best Local Similarity 65.6%; Pred. No. 1.7e-117;
Matches 210; Conservative 31; Mismatches 45; Indels 34; Gaps 5;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
DB 51 AFLASHLLIQSLFAFLEHRRKMKKSLETPI--KLNTVALCIAAYQEDPDYLRKCLQSVKR 108
QY 61 LLYPRARLVLMVVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQFWPEPAAAGAVGA 120
DB 109 LLYPRARLVLMVVDGNRAEDLYMVDMPREVFADDPATYVWDGNYHQFWPEPAAAGAVGA 120
QY 121 GAYREVEAEDPGLAVEALVTRRCVCVAQRWGGKREVMYTAFAKALGSDVDYVQVCDSDT 180
DB 158 GETEESHKESQH--VTQLVLSNKSICIMQKGGKREVMYTAFAKALGSDVDYVQVCDSDT 215
QY 181 RLDPMALLELVRLVLEDDPRVAGVGDVRLNPLDSWVSFLSLRYWVAFNVERACQSYFH 240
DB 216 RLDPMALLELVRLVLEDDPRVAGVGDVRLNPLDSWVSFLSLRYWVAFNVERACQSYFH 275
QY 241 CVSCISGPLGLYRNLLQOQFLEAWYKFLGTHCTFGDDRHLTNRMLSMGYATK----- 294
DB 276 CVSCISGPLGLYRNLLQOQFLEAWYKFLGTHCTFGDDRHLTNRMLSMGYATK----- 335
QY 295 -----AEGTRWS 301
DB 336 CYSETPSSFLRWLSQQTWS 355

Search completed: March 11, 2005, 14:17:16

Job time : 91 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 14:06:48 ; Search time 18.7879 Seconds
(without alignments)
1638.789 Million cell updates/sec

Title: US-10-672-399-8
Perfect score: 1697
Sequence: 1 AFLSAHLVAQSLFAYLEHRR.....SGTPGPAATRRPRPCGG 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	1550.5	91.4	543	2 JC4812	hyaluronan synthas
2	965.5	56.9	588	2 A43740	DG42 protein - Afr
3	409	24.1	568	2 T17588	hyaluronoglucosami
4	292	17.2	424	1 ZZRCL	nodulation protein
5	279	16.4	426	2 A95321	NodC N-ACETYLGLUCO
6	276	16.3	424	1 ZZRCL	nodulation protein
7	261	15.4	424	1 S12793	nodulation protein
8	257.5	15.2	413	1 S34305	nodulation protein
9	226.5	13.3	395	1 JQ3396	nodulation protein
10	214	12.6	366	2 E59102	hypothetical prote
11	213.5	12.6	395	2 A48755	hyaluronan synthas
12	213.5	12.6	419	2 A53100	hyaluronate syntha
13	185	10.9	1086	2 JC6079	chitin synthase (E
14	177	10.4	1195	2 S61886	chitin synthase (E
15	174	10.3	1165	1 S45879	chitin synthase (E
16	172	10.1	1112	2 T30202	probable chitin sy
17	168	9.9	1041	2 T31097	chitin synthase (E
18	167.5	9.9	1239	2 T42020	class IV chitin sy
19	160	9.4	417	2 JC4547	polysaccharide syn
20	158	9.3	428	1 E38180	nodulation protei
21	157	9.3	414	2 S18962	PBF15 protein - St
22	156.5	9.2	1175	2 S39951	chitin synthase (E
23	148.5	8.8	1103	2 T42022	probable chitin sy
24	143	8.4	1869	2 A59290	class V chitin syn
25	136	8.0	1498	2 S78102	chitin synthase (E
26	133	7.8	1852	2 JC5546	chitin synthase (E
27	132	7.8	447	2 A97211	glycosyltransferas
28	127	7.5	743	2 T34632	probable bi-functi
29	124	7.3	1380	2 T25284	hypothetical prote

30	119	7.0	482	2 G83928	hypothetical prote
31	117	6.9	420	2 D69769	cellulose synthase
32	113	6.7	1009	2 S20538	chitin synthase (E
33	112	6.6	420	2 AG2473	hypothetical prote
34	111.5	6.6	412	2 B90075	intercellular adhe
35	106	6.2	352	2 F75099	rhannosyl transfer
36	104.5	6.2	412	2 S77608	probable intercell
37	100.5	5.9	425	2 F97108	probable glycosylt
38	100	5.9	1013	2 JC2314	chitin synthase (E
39	99.5	5.9	1155	2 AC2675	chromosome segrega
40	99.5	5.9	1165	2 A97457	structural maintena
41	98	5.8	869	2 H83500	probable glucosyl
42	97.5	5.7	479	2 S75693	hypothetical prote
43	97	5.7	1774	2 T17421	polyketide synthas
44	96.5	5.7	914	2 S18942	hypothetical prote
45	95	5.6	963	2 S45167	chitin synthase (E

ALIGNMENTS

RESULT 1

JC4812
hyaluronan synthase (EC 2.4.1.1.-) - human
C:Species: Homo sapiens (man)
C>Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 02-Mar-2001
C:Accession: JC4812
R:Itano, N.; Kimata, K.
Biochem. Biophys. Res. Commun. 222, 816-820, 1996
A:Title: Molecular cloning of human hyaluronan synthase.
A:Reference number: JC4812, MUID:96244584; PMID:8651928
A:Accession: JC4812
A:Molecule type: mRNA
A:Residues: 1-543 <ITA>
A:Cross-references: DBJ:D84424; NID:G1401033; PIDN:BAAL2351.1; PID:d1013030; PID:g1401033
C:Comment: This enzyme synthesizes a high molecular weight glycosaminoglycan composed of
C:Genetics:
A:Gene: GDB:HAS1; HAS
A:Cross-references: GDB:1220109; OMIM:601463
A:Map position: 19q13.4-19q13.4
C:Keywords: glycosyltransferase; hexosyltransferase; phosphoprotein
F:58,167,308,321/Binding site: phosphate (Thr) (covalent) #status predicted
F:82,247/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match	91.4%	Score 1550.5;	DB 2;	Length 543;
Best Local Similarity	92.8%	Pred. No. 7.6e-128;		
Matches 297;	Conservative 3;	Mismatches 1;	Indels 19;	Gaps 1;
QY	1	AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA	60	
Db	26	AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA	85	
QY	61	LLYPRARLRLVMDGNRAEDLYVMDMPREVFADDPATYVWDGNYHQWPEAAAGAVGA	120	
Db	86	LLYPRARLRLVMDGNRAEDLYVMDMPREVFADDPATYVWDGNYHQWPEAAAGAVGA	145	
QY	121	GAYREVEADPGRLAVALVTRRCVCVQAQRWGGKREVMYTAFAKALGDSVDVQVCDSDT	180	
Db	146	GAYREVEADPGRLAVALVTRRCVCVQAQRWGGKREVMYTAFAKALGDSVDVQVCDSDT	205	
QY	181	RLDPMALLLEVRVLDDEPRVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSFH	240	
Db	206	RLDPMALLLEVRVLDDEPRVGAAGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSFH	265	
QY	241	CVSCISGPLGLYRNLLQOFLQFLGTHCTFGDDRHLTNRLMSGYATK-----	294	
Db	266	CVSCISGPLGLYRNLLQOFLQFLGTHCTFGDDRHLTNRLMSGYATKTSRSR	325	
QY	295	-----ABGTWS 301		
Db	326	CYSETPSSFLRWLSQQTRWS 345		

```
RESULT 2
A43740
D442 protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 09-Jul-2004
C:Accession: A43740
R:Roza, F.; Sargent, T.D.; Rebbert, M.L.; Michaelis, G.S.; Jamrich, M.; Grunz, H.; Jonas,
Dev. Biol. 129, 114-123, 1988
A>Title: Accumulation and decay of DG42 gene products follow a gradient pattern during x
A:Reference number: A43740; MUID:88313363; PMID:3410156
A:Accession: A43740
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-568 <ROS>
A:Cross-references: UNIPROT:P13563; EMBL:M22249; NID:g214098; PID:AAA49699.1; PID:g2140

Query Match 56.9%; Score 965.5; DB 2; Length 588;
Best Local Similarity 60.5%; Pred. No. 1.4e-76;
Matches 184; Conservative 47; Mismatches 62; Indels 11; Gaps 4;

QY 6 HLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYOEDPAYLROCLASARALLIYPR 65
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
72 HLMQSLFAFLERIV--NKSFLPCFKKTVALTIAGYQENPEYLILKLSCKTVKYPK 128
QY 66 ARLRLVMVVDGNRAEDLYMVDMPREVEADEDPATVYVMDGNVHPWEPAAAGAVGAGAYRE 125
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
129 DKLKIILVIDGNTEDDATMMFKDVFHGEDVGTIVWKGNTHTVKKPE---ETNKGSCPE 185
QY 126 VEA---EDPGLRLAVALVTRRCVCVAQRWGGKREVMYTFKALGDSVDYVQVCDSDTRL 182
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
186 VSKPLNEDEGINWVEELRNKRCVCIMQMGWGKREVMYTFQAIGTSVDYVQVCDSDTKL 245
QY 183 DPMALLELVRLDEDPVGVAGVDVRIINPLDSWVSFTSSRYWVAFVNERACQSYFHCV 242
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
246 DELATVEMVKVLESNDMYGAVGDVRIINPNYDSFISFMSSRLRYWVAFVNERACQSYFDCV 305
QY 243 SCISGLPLGYRNLLQOFLAWYNQKFLGTHCTFGDDRHLTNRLMSMGVATK--AEGTRW 300
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
306 SCISGLPLGYRNLLQVLEAWTRQKFLGTYCTLGGDRHLTNRLMSMGTRIKYTHKSRAP 365
QY 301 SGP 304
Db |||
366 SGP 369

RESULT 3
T17588
hyaluronoglucosaminidase-like protein - Chlorella virus PBCV-1
N:Alternate names: hyaluronic acid synthetase
C:Species: Chlorella virus PBCV-1
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T17588
R:Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A:Reference number: Z18806
A:Accession: T17588
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-568 <GRA>
A:Cross-references: UNIPROT:Q84419; EMBL:U42580; NID:G4028896; PIDN:AAC96466.1
C:Genetics:
A>Note: A98R

Query Match 24.1%; Score 409; DB 2; Length 568;
Best Local Similarity 32.9%; Pred. No. 7.4e-28;
Matches 103; Conservative 63; Mismatches 105; Indels 42; Gaps 10;

QY 2 FLSAHLVAQSLFAYLEHRR-----VAAARGPLDAATARSVALTISAYOEDPAYLROCLAS 57
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
56 FVFGFLAQVLFSELNRKLRKWIISLRPKGMNDV----RLAVIIAGYREDPYMFQKCLUES 111
QY 58 ARALLYPRARLVLMVDGNRAEDLYMVDMPREVEADEDPATVYVMDGNVHPWEPAAAGA 117
```

```
Db 112 VRSDYGNV-ARLICVIDGDEDDDMRMAAVYKAIYND-----NIKKP----- 152
QY 118 VGAGAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMYTFK--ALGDSVDYVQV 175
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
153 ----EFVLCESDDKEGERIDS--DFSRDICVLQPHRGKRECLYTGFOQLAKMDPVSNAVVL 206
QY 176 CDSSTRLDPMALLELVRLDEDPVGVAGVDVRIINPLDSWVSFLSSLRYWVAFVNERAC 235
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
207 IDSTVLEKDAILEVVVPLACDPFIQAVAGECKLWN-TDILLSLLVAWRYYSACVVERSA 265
QY 236 QSYFHCVCSCISGPLGYRNLLQOFLAWYNQKFLGTHCTFGDDRHLTNRLMSMG----Y 291
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
266 QSPFTVQCVGGPLGAYKIDIIKEIKDPWISQRFQKQCTYGGDRRLTNEILMRGKKVVF 325
QY 292 ATKAEGRWSTGP 304
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
326 TPFVAVG--WSDSP 336

RESULT 4
ZZZRCL
nodulation protein nodC - Rhizobium leguminosarum plasmid pRL1J1
C:Species: Rhizobium leguminosarum
C>Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C:Accession: A03486
R:Rozen, L.; Johnston, A.W.B.; Downie, J.A.
Nucleic Acids Res. 12, 9497-9508, 1984
A>Title: DNA sequence of the Rhizobium leguminosarum nodulation genes nodAB and C requir
A:Reference number: A03482; MUID:85087952; PMID:6514582
A:Accession: A03486
A:Molecule type: DNA
A:Residues: 1-424 <ROS>
A:Cross-references: UNIPROT:P04340; GB:X01650; NID:g46212; PIDN:CAA68619.1; PID:g46215
C:Comment: This is one of the proteins, coded by nodulation genes, that are required for
C:Genetics:
A:Gene: nodC
A:Genome: plasmid
A:Superfamily: nodulation protein nodC
C:Keywords: nodulation

Query Match 17.2%; Score 292; DB 1; Length 424;
Best Local Similarity 26.6%; Pred. No. 9.1e-18;
Matches 84; Conservative 47; Mismatches 113; Indels 72; Gaps 8;

QY 7 LVAQSLFAYLE-----HRRVAAARGP---LDAATARSVALTISAYOEDPAYLRQC 54
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
9 IAAISLYAMLSTVYKSAQVFHARETTISTTPAKDIETNPVPSVDVIVPCFNEEDPTVLSEC 68
QY 55 LASARALLYPRARLVLMVDGNRAEDLYMVDMPREVEADEDPATVYVMDGNVHPWEPAA 114
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
69 LASLAEQDY-AGKURIVVDGSKNRDAVVAQ--RAAYADDERENFT----- 112
QY 115 AGAVGAGAYREVEADEPGRLAVALVTRRCVCVAQRWGGKREVMYTFKALGDSVDYVQ 174
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
113 -----ILPKNVGKRKAIAAITQSSGDLILNV- 138
QY 175 VCDSDTRLDPWALLELVRLDEDPVGVAGVDVRIINPLDSWVSFLSSLRYWVAFVNERA 234
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
139 --DSDTTIAPDVVSKLAHKM-RDPAVGAAMQMKASNAQADTWLTRLIDMEYWLACNEERA 195
QY 235 QOSYFHCVCSCISGPLGYRNLLQOFLAWYNQKFLGTHCTFGDDRHLTNRLMSMGVATK 294
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
196 AARFPGAVMCCGPGCMYRRSAMLSDIDQYETQLYRGKPSDFGEDRHLTILMSLGSFRTE 255
QY 295 AEGTRWSTGTPGPAAT 310
Db |||
256 Y-----VPSAIAAT 264

RESULT 5
A95321
NodC N-ACETYLGLUCOSAMINYLTRANSFERASE [imported] - Sinorhizobium meliloti (strain 1021) m
C:Species: Sinorhizobium meliloti
```


Db 3 VAVVPSYNESASAIWNTINSVLAQDYPIH--EIPFVDDGSKDKAGAYEVALMKRRELLRT 60
QY 95 EDPATYVWDGNYHQWPAPAAAGVAGAGAYREVEAEDPGRGLAVEALVTRRCVC----- 147
Db 61 Q-----REIAA-----TTKNICSEILGIP 79
QY 148 --VAQRWG---CKREWYTFKALGSDVDYVQVCDSTRLDPMALLELVRLVDEDPVGA 202
Db 80 DLIVHRLPKNCGRKRAQLWAFR--TTDAIVTIDSDGLFFNAVRELLKPFN-DEKVA 136
QY 203 VCGDVRILNPLDSWSYSSLSYVAFNVERACQSYFHCVSCISGPGLYRNLLAQOFL 262
Db 137 TTGHVNRNRNDNLTKLIDMDYDNFRVRAAQSYVTGNLVCSGPLSYRREVTENLE 196
QY 263 AWYQKFLGTHCTFGDDRLHTRNRLSMG 290
Db 197 HYGSMFLGEEVQFGDDRCCLTNVAILKG 224
RESULT 11
A48755
hyaluronan synthase (EC 2.4.1.-) - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
C>Date: 19-May-1995 #sequence_revision 19-May-1995 #text_change 09-Jul-2004
C:Accession: A48755; JC2077
R:DeAngelis, P.L.; Papaconstantinou, J.; Weigel, P.H.
J. Biol. Chem. 268, 19181-19184, 1993
A:Title: Molecular cloning, identification, and sequence of the hyaluronan synthase gene
A:Reference number: A48755; MUID:93374890; PMID:8366070
A:Accession: A48755
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-395 <DEA1>
A:Cross-references: UNIPROT:Q54866; UNIPROT:Q54867; UNIPROT:Q8K5G6; GB:L20853; NID:94100
R:De Angelis, P.L.; Yang, N.; Weigel, P.H.
Biochem. Biophys. Res. Commun. 199, 1-10, 1994
A:Title: The Streptococcus pyogenes hyaluronan synthase: Sequence comparison and conserv
A:Reference number: JC2077; MUID:94168559; PMID:8122999
A:Accession: JC2077
A:Molecule type: DNA
A:Residues: 1-395 <DEA2>
C:Comment: This enzyme is responsible for polymerizing an alternating polysaccharide cop
C:Genetics:
A:Gene: hasA
C:Superfamily: modulation protein nodC
C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein
F:9-29/Domain: transmembrane #status predicted <TM1>
F:294-314/Domain: transmembrane #status predicted <TM2>
F:328-341/Domain: transmembrane #status predicted <TM3>
F:353-370/Domain: transmembrane #status predicted <TM4>
Query Match 12.6%; Score 213.5; DB 2; Length 395;
Best Local Similarity 28.1%; Pred. No. 6.3e-11; Mismatches 27; Indels 67; Gaps 6;
Matches 74; Conservative 27; Mismatches 95; Indels 67; Gaps 6;
QY 37 VALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAEDLYM-----VDM 87
Db 42 VAAVPSYNEDEAESLETLKSVLAQTYPLS--EIIVDDGSSNTDAIQLEIEVYVNRVDI 99
QY 88 FREVPADDDPATYVWDGNYHQWPAPAAAGVAGAGAYREVEADPGRGLAVEALVTRRCVC 147
Db 100 CRNVIVHR---SLVNKGRKH----- 116
QY 148 VAQRWGGKREVMYTAFAKALGSDVDYVQVCDSTRLDPMALLELVRLVDEDPVGAAGVDV 207
Db 117 -AQAWAFER-----SDADVLTVDSDTYIYPNALLELLKSFN-DETVYAATGHL 163
QY 208 RILNPLDSWSYSSLSRYVAFNVERACQSYFHCVSCISGPGLYRNLLAQOFL 267
Db 164 NARNRQTLLTLRLTDIRDYDNAGVERAAQSLTGNILVCSGPLSIYRREVIIPNLERYKNQ 223
QY 268 KFLGTHCTFGDDRLHTRNRLSMG 290

Db 224 TFLGLPVSIGDDRCCLTNVAILDG 246
RESULT 12
A53100
hyaluronate synthase A (HasA) - Streptococcus sp. (group A)
C:Species: Streptococcus sp.
C>Date: 06-Oct-1994 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C:Accession: A53100
R:Dougherty, B.A.; van de Rijn, I.
J. Biol. Chem. 269, 169-175, 1994
A:Title: Molecular characterization of hasA from an operon required for hyaluronic acid
A:Reference number: A53100; MUID:94103204; PMID:8276791
A:Accession: A53100
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-419 <DOU>
A:Cross-references: UNIPROT:Q8NKK1
A>Note: sequence extracted from NCBI backbone (NCBIN:141683, NCBI:141684)
C:Superfamily: modulation protein nodC
Query Match 12.6%; Score 213.5; DB 2; Length 419;
Best Local Similarity 28.1%; Pred. No. 6.7e-11; Mismatches 27; Indels 67; Gaps 6;
Matches 74; Conservative 27; Mismatches 95; Indels 67; Gaps 6;
QY 37 VALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDGNRAEDLYM-----VDM 87
Db 66 VAAVPSYNEDEAESLETLKSVLAQTYPLS--EIIVDDGSSNTDAIQLEIEVYVNRVDI 123
QY 88 FREVPADDDPATYVWDGNYHQWPAPAAAGVAGAGAYREVEADPGRGLAVEALVTRRCVC 147
Db 124 CRNVIVHR---SLVNKGRKH----- 140
QY 148 VAQRWGGKREVMYTAFAKALGSDVDYVQVCDSTRLDPMALLELVRLVDEDPVGAAGVDV 207
Db 141 -AQAWAFER-----SDADVLTVDSDTYIYPNALLELLKSFN-DETVYAATGHL 187
QY 208 RILNPLDSWSYSSLSRYVAFNVERACQSYFHCVSCISGPGLYRNLLAQOFL 267
Db 188 NARNRQTLLTLRLTDIRDYDNAGVERAAQSLTGNILVCSGPLSIYRREVIIPNLERYKNQ 247
QY 268 KFLGTHCTFGDDRLHTRNRLSMG 290
Db 248 TFLGLPVSIGDDRCCLTNVAILDG 270
RESULT 13
JC6079
chitin synthase (EC 2.4.1.16) chsD - Emericella nidulans
C:Species: Emericella nidulans, Aspergillus nidulans
C>Date: 21-Jan-1997 #sequence_revision 21-Jan-1997 #text_change 20-Jun-2000
C:Accession: JC6079
R:Moroyama, T.; Fujiwara, M.; Kojima, N.; Horiuchi, H.; Ohta, A.; Takagi, M.
Mol. Gen. Genet. 251, 442-450, 1996
A:Title: The Aspergillus nidulans genes chsA and chsD encode chitin synthases which have
A:Reference number: JC6079; MUID:96285568; PMID:8709948
A:Accession: JC6079
A:Molecule type: DNA
A:Residues: 1-1086 <MOT>
A:Cross-references: DDBJ:D83246; NID:gl688025; PIDN:BAAL1866.1; PID:gl688026
C:Comment: This enzyme functions in conidia formation.
C:Genetics:
A:Gene: chsD
A:Introns: 800/2
C:Superfamily: chitin synthase chs4
C:Keywords: glycosyltransferase; hexosyltransferase
Query Match 10.9%; Score 185; DB 2; Length 1086;
Best Local Similarity 22.0%; Pred. No. 6.7e-08; Mismatches 54; Indels 56; Gaps 11;
Matches 71; Conservative 54; Mismatches 131; Indels 56; Gaps 11;
QY 28 PLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLRVLMVVDG---NRAEDLYM 84

Db 563 PYGPPLAHACL-VTCYSEGEIGRTTLDLSIALTDYPNSHKSIIVVICDGIILKKGEEFST 621
Qy 85 VDMFREVAD-----ED-----PATYVDGNVHQHPWEPAAAGAGAGAYREVEADPGRLA 135
Db 622 PDIVLMMRDPIIPPEEYEAFFYAVATGSKRHNNMAKYAGFYDGEHSIIIPVENMQQRVP 681
Qy 136 VEALVRRTRCVCVAQRWG-----GKR-----EVMYTAFAKAL--- 166
Db 682 MMIVV---KCGTPAATAAKPQNGRGRDSQIILMSFLQKVPFDERMTELEYEMFNGLLHV 738
Qy 167 ----GDSVDYVQVCDSTRDPMALLELVRLVDEDPYRGVAGDGVRIILNPLDLSWVSFLSS 222
Db 739 TGIPDPDFEVLVMDADTKVPFDSLTHMISAMVKDPEVMGLCGETKIANKTDSWVTMIQV 798
Qy 223 LRYWAFNVERACOSYFHCVCISGLPLGLYRNLLQQLFLEAW-----YNQKFL 270
Db 799 FEYFVSHHAKAFESVFGVTCFPCFMYRIKAPKQGNQYVWPILANPDVIVHSENVV 858
Qy 271 GT-H-----CTFGDDRHLNRMML 287
Db 859 DTLHKNNLLLGEDRYLSTLML 880

RESULT 14

S61886

chitin synthase (EC 2.4.1.16) CHS4 - Neurospora crassa

N:Alternate names: chitin synthase class IV

C:Species: Neurospora crassa

C:Date: 23-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004

C:Accession: S61886

R:Dir, A.B.; Specht, C.A.; Robbins, P.W.; Yarden, O.

Mol. Gen. Genet. 250, 214-222, 1996

A:Title: chs-4, a class IV chitin synthase gene from Neurospora crassa.

A:Reference number: S61886; MUID:96188842; PMID:8628221

A:Accession: S61886

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-1195 <DIN>

A:Cross-references: UNIPROT:Q01285; EMBL:U25097; NID:G793937; PIDN:AAB03563.1; PID:G7939

C:Genetics:

A:Gene: chs-4

A:Introns: 1042/1

C:Superfamily: chitin synthase chs4

C:Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase; transmembrane

Query Match 10.4%; Score 177; DB 2; Length 1195;

Best Local Similarity 22.4%; Pred. No. 3.8e-07;

Matches 72; Conservative 54; Mismatches 130; Indels 66; Gaps 12;

Qy 28 PLDAATARSVALTISAYQEDPAYLRQCLASARALLYPRARLVLMVVDG---NRAEDLYM 84
Db 641 PFGPPLAHTICL-VTAYSEGMGVRITLDSLTAMTDYPNSHKVILVICDGIILKKGHEHST 699
Qy 85 VDMFREVAD-----DPATY--VMDGNVHQHPWEPAAAGAGAGAYREVEADPGRLA 135
Db 700 PDILGHMKDHTIHPDDVEPFSYAVATGSKRHNNMAKYTFYDGTNSAIPLEKQQRVP 759
Qy 136 VEALVRRTRCVCVAQRW-----GKR-----EVMYTAFAKAL--- 166
Db 760 MMVVV---KCGTPAEASKPKNGRGRDSQIILMSFLQKVPFDERMTELEYEMFNGLLWKI 816
Qy 167 -GDSVDYVQV---CDSVTRLDPMALLELVRLVDEDPYRGVAGDGVRIILNPLDLSWVSFLSS 222
Db 817 TGISPDPFYEIVLMDADTKVPFDSLTHMISAMVKDPEIMGLCGETKIANKRAASWVSAIQV 876
Qy 223 LRYWAFNVERACOSYFHCVCISGLPLGLYRNLLQQLFLEAW-----YNQKFL 270
Db 877 FEYFVSHHAKAFESVFGVTCFPCFMYRIKAPKQGNQYVWPILANPDVIVHSENVV 936
Qy 271 GT-H-----CTFGDDRHLNRMML 287
Db 937 DTLHKNNLLLGEDRYLSTLML 958

RESULT 15

S45879

chitin synthase (EC 2.4.1.16) 3 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein YBR023c; protein YBR0305

C:Species: Saccharomyces cerevisiae

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: S45879; S46554; S22776; A39639; S17247

R:Grivell, L.A.; de Haan, M.; Maat, M.J.; Smits, P.H.M.

submitted to the Protein Sequence Database, August 1994

A:Reference number: S45875

A:Accession: S45879

A:Molecule type: DNA

A:Residues: 1-1165 <GRI>

A:Cross-references: UNIPROT:P29465; EMBL:Z35892; NID:G536229; PIDN:CAA84965.1; PID:G5362

A:Experimental source: strain S288C

R:Smits, P.H.M.; de Haan, M.; Maat, C.; Grivell, L.A.

Yeast 10(Suppl.A), S75-S80, 1994

A:Title: The complete sequence of a 33 kb fragment on the right arm of chromosome II fro

ly identified genes and a homologue of the SCO1 gene.

A:Reference number: S46551; MUID:94378725; PMID:8091864

A:Accession: S46554

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1165 <SMI>

A:Cross-references: EMBL:X76078; NID:G498748; PIDN:CAAS3680.1; PID:G498752

A:Experimental source: strain S288C

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1993

R:Bulawa, C.E.

Mol. Cell. Biol. 12, 1764-1776, 1992

A:Title: CSD2, CSD3, and CSD4, genes required for chitin synthesis in Saccharomyces cere

bium species and Xenopus laevis.

A:Reference number: S22776; MUID:92195323; PMID:1532231

A:Accession: S22776

A:Molecule type: DNA

A:Residues: 1-1162,'L',1164-1165 <BUL>

A:Cross-references: EMBL:M73697; NID:G172103; PIDN:AAA34844.1; PID:G172104

R:Valdivieso, M.H.; Mol, P.C.; Shaw, J.A.; Cabib, E.; Duran, A.

J. Cell Biol. 114, 101-109, 1991

A:Title: CAL1, a gene required for activity of chitin synthase 3 in Saccharomyces cerevi

A:Reference number: A39639; MUID:91268144; PMID:2050737

A:Accession: A39639

A:Molecule type: DNA

A:Residues: 67-1165 <VAL>

A:Cross-references: GB:X57300; NID:G3359; PIDN:CAA40559.1; PID:G3360

C:Genetics:

A:Gene: SGD:CHS3; CAL1; CSD2; MIPS:YBR023c

A:Cross-references: MIPS:YBR023c; SGD:S0000227

A:Map position: 2R

C:Function:

A:Description: catalyzes the alpha-1,4-glycosylation of chitin by UDP-N-acetyl-D-glucosa

C:Superfamily: chitin synthase chs4

C:Keywords: chitin biosynthesis; glycosyltransferase; hexosyltransferase; transmembrane

F:203-219/Domain: transmembrane #status predicted <TM1>

F:457-473/Domain: transmembrane #status predicted <TM2>

F:1018-1034/Domain: transmembrane #status predicted <TM3>

F:1035-1054/Domain: transmembrane #status predicted <TM4>

F:1060-1076/Domain: transmembrane #status predicted <TM5>

F:1084-1100/Domain: transmembrane #status predicted <TM6>

Query Match 10.3%; Score 174; DB 1; Length 1165;

Best Local Similarity 21.1%; Pred. No. 6.7e-07;

Matches 67; Conservative 59; Mismatches 108; Indels 84; Gaps 10;

Qy 41 ISAYQEDPAYLRQCLASARALLYPRARLVLMVVDG---NRAED-----LYMVDMPRE 90
Db 658 VTCYSEDEGLRTTLDLSLTDDYPNSHKLMVWCDGLIKSGNDKTTPEIAGMWDFFVT 717
Qy 91 VFADEDPATYVDGNVHQHPWEPAAAGA-----VGAGAYR-----EVEADPGRLAVAL 139
Db 718 PPDEVKPYSYV-----AVASGSKRHNNMAKIYAGFYKDDSTIPPENQQRVPIT 767
Qy 140 VRTRRCVCAQRWG-----GKR-----EVMYTAFA-----KAL 166

Db 768 V---KCGTPAEQGAAPGNRGRDSQIILMSFLEKITFDERMTQLEFOLLKNIWOITGLM 824
Qy 167 GDSVDYVOVCDSDTRLDPMALLELVRLDEDPVGAAGDVRIINPLDSWVSFLSSLYW 226
Db 825 ADPYETVLMVDADTKVFPDALTHVAEMVKDPLIMGLCGETKIANKAQSWTAIOVFYY 884
Qy 227 VAFNVERACQSYFHCVCISGELGLYR-----NNLLQQFLEAWYNQKF 269
Db 885 ISHQAKAFESVFGSVTCLPGCFSMYRIKSPKSGDYWVFLANPDIVERYSNDVNTLH 944
Qy 270 LGTHCTFGDDRHLITNRL 287
Db 945 KKNLLLGEDRFLSSML 962

Search completed: March 11, 2005, 14:23:14
Job time : 19.7879 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 13:48:17 ; Search time 83.6364 Seconds
(without alignments)
1959.259 Million cell updates/sec

Title: US-10-672-399-8

Perfect score: 1697

Sequence: 1 AFLSAHLVAQSLFAYLEHRR.....SGTPPGPAATQRRPRPCGG 320

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_prot.*
2: uniprot_trebl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1553.5	91.5	578	HAS1_HUMAN	Q92839 homo sapien
2	1553.5	91.5	578	Q9NS49	Q9ns49 homo sapien
3	1550.5	91.4	577	Q81YH3	Q81yh3 homo sapien
4	1508.5	88.9	583	HAS1_MOUSE	Q61647 mus musculus
5	1508	88.9	583	Q6S742	Q6s742 papio anubi
6	1505.5	88.7	583	Q8CH93	Q8ch93 rattus norv
7	1048.5	61.8	458	Q67488	Q6t488 brachydanio
8	965.5	56.9	588	HAS1_XENLA	P13563 xenopus lae
9	948.5	55.9	552	Q95M29	Q95m29 oryctolagus
10	938	55.3	553	Q75R37	Q75r37 sus scrofa
11	937	55.2	552	Q8SQ70	Q8sq70 sus scrofa
12	935.5	55.1	554	HAS3_MOUSE	Q08650 mus musculus
13	935	55.1	552	HAS2_MOUSE	P70312 mus musculus
14	935	55.1	552	Q8HZJ3	Q8hzj3 equus caball
15	935	55.1	552	Q95M16	Q95m16 oryctolagus
16	934	55.0	552	HAS2_HUMAN	Q92819 homo sapien
17	933.5	55.0	554	Q8CEB9	Q8ceb9 mus musculus
18	933.5	55.0	554	Q8CH92	Q8ch92 rattus norv
19	933	55.0	552	HAS2_BOVIN	Q97711 bos taurus
20	930	54.8	552	HAS2_RAT	Q35776 rattus norv
21	928.5	54.7	554	Q8DGD0	Q9gd40 brachydanio
22	928	54.7	553	HAS3_HUMAN	Q00219 homo sapien
23	928	54.7	553	Q96RV2	Q96rv2 homo sapien
24	927	54.6	557	Q8W9J2	Q8w9j2 xenopus lae
25	924	54.4	552	Q9DGA1	Q9dga1 brachydanio
26	923	54.4	552	HAS2_CHICK	Q57424 gallus galli
27	907	53.4	551	HAS2_XENLA	O57427 xenopus lae
28	843.5	49.7	583	HAS3_XENLA	O57428 xenopus lae
29	842.5	49.6	583	Q8AZI0	Q6azi0 xenopus lae
30	742.5	43.8	393	Q811Y6	Q81ly6 rattus norv
31	552.5	32.6	156	Q18792	O18792 papio anubi

32	528.5	31.1	281	2	Q8WTZ0	Q8wtz0 homo sapien
33	488	28.8	134	2	Q90489	Q90489 brachydanio
34	473.5	27.9	245	2	Q9GK14	Q9gk14 bos taurus
35	409	24.1	568	2	Q84419	Q84419 paramesicium
36	409	24.1	568	2	Q9WFS9	Q9wfs9 paramesicium
37	409	24.1	568	2	Q9WFT0	Q9wft0 paramesicium
38	355.5	20.9	172	2	Q9XS25	Q9xs25 bos taurus
39	301.5	17.8	131	2	Q8BPN0	Q8bpn0 mus musculus
40	292	17.2	424	1	NODC_RHILV	P04340 rhizobium l
41	290.5	17.1	452	2	Q9AQZ3	Q9aqz3 bradyrhizob
42	279.5	16.5	443	2	Q8KLG3	Q8klg3 rhizobium e
43	279	16.4	426	1	NODC_RHIME	P04341 rhizobium m
44	276.5	16.3	452	2	Q6EX51	Q6ex51 sinorhizobi
45	274	16.1	402	2	O52478	O52478 rhizobium m

ALIGNMENTS

RESULT 1
HAS1_HUMAN
ID HAS1_HUMAN STANDARD; PRT; 578 AA.
AC Q92839; Q14470;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE (Hyaluronic acid synthase 1) (HA synthase 1) (HuHAS1).
GN Name=HAS1; Synonyms=HAS;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymph node;
RX MEDLINE=96394438; PubMed=8798544; DOI=10.1074/jbc.271.38.23395;
RA Shyjan A.M., Heidlin P., Butcher E.C., Yoshino T., Brisken M.J.;
RT "Functional cloning of the cDNA for a human hyaluronan synthase.";
RL J. Biol. Chem. 271:23395-23399(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=96244584; PubMed=8651928; DOI=10.1006/bbrc.1996.0827;
RT Itano N., Kinata K.;
RT "Molecular cloning of human hyaluronan synthase.";
RL Biochem. Biophys. Res. Commun. 222:816-820(1996).
CC -!- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA) synthesis. Also able to catalyze the synthesis of chito-oligosaccharide depending on the substrate.
CC -!- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-glucuronosyl(1->3)](n) + 2n UDP.
CC -!- Cofactor: Magnesium.
CC -!- PATHWAY: Hyaluronate synthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- TISSUE SPECIFICITY: Highly expressed in ovary followed by spleen, thymus, prostate, testes and large intestine. Weakly expressed in small intestine.
CC -!- SIMILARITY: Belongs to the nodC/HAS family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; U59269; AAC50706.1; -;
EMBL; D84424; BAA12351.1; ALT_INIT.
Genew; HGNC:4818; HAS1.
MIN; 601463; -.

```

DR GO: 0005887; C: integral to plasma membrane; TAS.
DR GO: 0007155; P: cell adhesion; TAS.
DR GO: 0006024; P: glycosaminoglycan biosynthesis; TAS.
DR InterPro; IPR001173; Glyco_trans_2.
DR Pfam; PF00535; Glycosyltransferase; Multigene family; Transferase; Transmembrane.
KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
FT DOMAIN 1 25 Cytoplasmic (Potential).
FT TRANSMEM 26 46 1 (Potential).
FT DOMAIN 47 52 Extracellular (Potential).
FT TRANSMEM 53 73 2 (Potential).
FT DOMAIN 74 399 Cytoplasmic (Potential).
FT TRANSMEM 400 420 3 (Potential).
FT DOMAIN 421 430 Extracellular (Potential).
FT TRANSMEM 431 451 4 (Potential).
FT DOMAIN 452 457 Cytoplasmic (Potential).
FT TRANSMEM 458 478 5 (Potential).
FT DOMAIN 479 497 Extracellular (Potential).
FT TRANSMEM 498 518 6 (Potential).
FT DOMAIN 519 540 Cytoplasmic (Potential).
FT TRANSMEM 541 561 7 (Potential).
FT DOMAIN 562 578 Extracellular (Potential).
FT CONFLICT 1 2 MR -> RS (in Ref. 2).
FT CONFLICT 34 34 G -> A (in Ref. 2).
SQ SEQUENCE 578 AA; 64884 MW; 355FD54B0899E43C CRC64;

Query Match 91.5%; Score 1553.5; DB 1; Length 578;
Best Local Similarity 93.1%; Pred. No. 6.2e-125;
Matches 298; Conservative 2; Mismatches 1; Indels 19; Gaps 1;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCCLASARA 60
Db 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCCLASARA 120

Qy 61 LLYPRARLRVLWVDGNGRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPAAGAVGA 120
Db 121 LLYPRARLRVLWVDGNGRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPAAGAVGA 180

Qy 121 GAYREVEADPGRLAVEALVTRRCVCVAQWGWGKREVMYTFKALGDSVDYVQVCDSDT 180
Db 181 GAYREVEADPGRLAVEALVTRRCVCVAQWGWGKREVMYTFKALGDSVDYVQVCDSDT 240

Qy 181 RLDPMALLEVRVLDEDPVGAAGDVRILNPLDSWSFLSSLRVYAFNVERACQSYFH 240
Db 241 RLDPMALLEVRVLDEDPVGAAGDVRILNPLDSWSFLSSLRVYAFNVERACQSYFH 300

Qy 241 CVSCISGPLGYRNLLQQFLFLEAWYKQFLGTHCTFGDDRHLTNRMLSMGYATK----- 294
Db 301 CVSCISGPLGYRNLLQQFLFLEAWYKQFLGTHCTFGDDRHLTNRMLSMGYATKTSRSR 360

Qy 295 -----AEGTRWS 301
Db 361 CYSETPSSFLRWLSQOTRWS 380

RESULT 2
Q9NS49 PRELIMINARY; PRT; 578 AA.
AC Q9NS49;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hyaluronan synthase 1.
GN Name=HAS1;
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E.; Krcmar J.P.; Skowronski E.; Viswanathan V.;
RA Burkhardt-Schultz K.; Gordon L.; Dias J.; Ramirez M.; Stilwagen S.;
RA Phan H.; Velasco N.; Do L.; Regala W.; Terry A.; Brower A.; Garnes J.;
RA Danganan L.; Erler A.; Christensen M.; Georgescu A.; Avila J.; Liu S.,
Andreise T., Trankheim M., Attix C., Amico-Keller G., Coefield J.,
Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,
Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A.,
Olsen A.S., Carrano A.V.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA DOE Joint Genome Institute;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC018755; AAF87845.1; -.
DR InterPro; IPR001173; Glyco_trans_2.
DR InterPro; IPR00276; GPCR_Rhodopsin.
DR InterPro; IPR002057; Isopen N synth.
DR PROSITE; PS00237; G PROTEIN RECEPTOR FL 1; UNKNOWN_1.
DR PROSITE; PS00185; IPNS 1; UNKNOWN_1.
SQ SEQUENCE 578 AA; 64831 MW; 2FE3A44B0D5380FF CRC64;

Query Match 91.5%; Score 1553.5; DB 2; Length 578;
Best Local Similarity 93.1%; Pred. No. 6.2e-125;
Matches 298; Conservative 2; Mismatches 1; Indels 19; Gaps 1;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCCLASARA 60
Db 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCCLASARA 120

Qy 61 LLYPRARLRVLWVDGNGRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPAAGAVGA 120
Db 121 LLYPRARLRVLWVDGNGRAEDLYWDMFREVFADEDPATYVWDGNYHQWPEPAAGAVGA 180

Qy 121 GAYREVEADPGRLAVEALVTRRCVCVAQWGWGKREVMYTFKALGDSVDYVQVCDSDT 180
Db 181 GAYREVEADPGRLAVEALVTRRCVCVAQWGWGKREVMYTFKALGDSVDYVQVCDSDT 240

Qy 181 RLDPMALLEVRVLDEDPVGAAGDVRILNPLDSWSFLSSLRVYAFNVERACQSYFH 240
Db 241 RLDPMALLEVRVLDEDPVGAAGDVRILNPLDSWSFLSSLRVYAFNVERACQSYFH 300

Qy 241 CVSCISGPLGYRNLLQQFLFLEAWYKQFLGTHCTFGDDRHLTNRMLSMGYATK----- 294
Db 301 CVSCISGPLGYRNLLQQFLFLEAWYKQFLGTHCTFGDDRHLTNRMLSMGYATKTSRSR 360

Qy 295 -----AEGTRWS 301
Db 361 CYSETPSSFLRWLSQOTRWS 380

RESULT 3
Q8IYH3 PRELIMINARY; PRT; 577 AA.
AC Q8IYH3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hyaluronan synthase 1.
GN Name=HAS1;
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.L.; Feingold E.A.; Grouse L.H.; Derge J.G.;
RA Klausner R.D.; Collins F.S.; Wagner L.; Shenmen C.M.; Schuler G.D.;
RA Altschul S.F.; Zeeberg B.; Buetow K.H.; Schaefer C.F.; Bhat N.K.;
RA Hopkins R.F.; Jordan H.; Moore T.; Max S.I.; Wang J.; Heide F.;
RA Datchenko L.; Marusina K.; Farmer A.A.; Rubin G.M.; Hong L.;
RA Stapleton M.; Soares M.B.; Bonaldo M.F.; Casavant T.L.; Scheetz T.E.,

```

RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywicki M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN
RC SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC035837; AAH35837.1; -
DR InterPro; IPR001173; Glyco trans 2.
DR InterPro; IPR00276; GPCR Rhodopsin.
DR InterPro; IPR002057; Isopen N synth.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS00185; IPNS_1; UNKNOWN_1.
SQ SEQUENCE 577 AA; 64788 MW; 61C9AC33AEDBD183 CRC64;
Query Match 91.4%; Score 1550.5; DB 2; Length 577;
Best Local Similarity 92.8%; Pred. No. 1.1e-124; Indels 19; Gaps 1;
Matches 297; Conservative 3; Mismatches 1;
QY 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 60 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAAATARSVALTISAYQEDPAYLRQCLASARA 119
QY 61 LLYPARLRLVMDGNRAEDLYMDMFREVPADEDPATYWDGNYHGFQWEPAAAGAVGA 120
Db 120 LLYPARLRLVMDGNRAEDLYMDMFREVPADEDPATYWDGNYHGFQWEPAAAGAVGA 179
QY 121 GAYREVEADPCRLAEALVTRRCVCVAQRGGKREVMYTAFAKALGDSVDVQVCDSDT 180
Db 180 GAYREVEADPCRLAEALVTRRCVCVAQRGGKREVMYTAFAKALGDSVDVQVCDSDT 239
QY 181 RLDPMLLELVRVLDEDPVAGVGDVRLNPLDSWVSFLSLRYWAFNVERACQSYFH 240
Db 240 RLDPMLLELVRVLDEDPVAGVGDVRLNPLDSWVSFLSLRYWAFNVERACQSYFH 299
QY 241 CVSCISGPIGLYRNLLQFLBANTYQKFLGTHCTFGDDRHLTNRLMSGVATK----- 294
Db 300 CVSCISGPIGLYRNLLQFLBANTYQKFLGTHCTFGDDRHLTNRLMSGVATKYTSRSR 359
QY 295 -----ASGTRWS 301
Db 360 CVSETPSSFLRWLSQOTRWS 379
RESULT 4
ID_HASI_MOUSE STANDARD; PRT; 583 AA.
AC Q61647;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
DE (Hyaluronic acid synthase 1) (HA synthase 1).
GN Name=Hasi; Synonyms=Has;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96215261; PubMed=8626618; DOI=10.1074/jbc.271.17.9875;

RA Itano N., Kimata K.;
RT "Expression cloning and molecular characterization of HAS protein, a
RT eukaryotic hyaluronan synthase.";
RL J. Biol. Chem. 271:9875-9878(1996).
RN [2]
RP MUTAGENESIS.
RX MEDLINE=20085071; PubMed=10617644; DOI=10.1074/jbc.275.1.497;
RA Yoshida M., Itano N., Yamada Y., Kimata K.;
RT "In vitro synthesis of hyaluronan by a single protein derived from
RT mouse HAS1 gene and characterization of amino acid residues essential
RT for the activity.";
RL J. Biol. Chem. 275:497-506(2000).
CC -!- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA)
CC synthesis. Also able to catalyze the synthesis of chito-
CC oligosaccharide depending on the substrate.
CC -!- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-
CC glucuronate = [beta-D-N-acetyl-D-glucosaminyl(1->4)beta-D-
CC glucuronosyl(1->3)](n) + 2n UDP.
CC -!- COFACTOR: Magnesium.
CC -!- PATHWAY: Hyaluronate synthesis.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: Belongs to the nodC/HAS family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D82964; BAAL1654.1; -
DR MGD; MGI:106590; Has1.
DR InterPro; IPR001173; Glyco trans 2.
DR Pfam; PF00535; Glycosyltransf_2; 1.
KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
FT DOMAIN 1 24 Cytoplasmic (Potential).
FT TRANSMEM 25 45 1 (Potential).
FT DOMAIN 46 51 Extracellular (Potential).
FT TRANSMEM 52 72 2 (Potential).
FT DOMAIN 73 404 Cytoplasmic (Potential).
FT TRANSMEM 405 425 3 (Potential).
FT DOMAIN 426 435 Extracellular (Potential).
FT TRANSMEM 436 456 4 (Potential).
FT DOMAIN 457 462 Cytoplasmic (Potential).
FT TRANSMEM 463 483 5 (Potential).
FT DOMAIN 484 501 Extracellular (Potential).
FT TRANSMEM 502 522 6 (Potential).
FT DOMAIN 523 545 Cytoplasmic (Potential).
FT TRANSMEM 546 566 7 (Potential).
FT DOMAIN 567 583 Extracellular (Potential).
FT TRANSMEM 584 84 Poly-Ala.
FT DOMAIN 85 519 Poly-Leu.
FT CARBOHYD D-2-E: Loss of both activities.
FT MUTAGEN 242 242 S->E: Loss of both activities.
FT MUTAGEN 311 311 S->N: No effect.
FT MUTAGEN 312 312 G->P: No effect.
FT MUTAGEN 313 313 P->G: No effect.
FT MUTAGEN 314 314 L->V: Loss of HA activity.
FT MUTAGEN 314 314 L->I: 75% decrease of both activities.
FT MUTAGEN 344 344 D->E: Loss of both activities.
FT MUTAGEN 380 380 Q->N: 85%-90% decrease of both activities.
FT MUTAGEN 383 383 R->K: 85%-90% decrease of both activities.
FT MUTAGEN 384 384 W->Y: Loss of both activities.
SQ SEQUENCE 583 AA; 65544 MW; 7AF9273E7B314728 CRC64;
Query Match 88.9%; Score 1508.5; DB 1; Length 583;
Best Local Similarity 89.9%; Pred. No. 4.7e-121; Indels 25; Gaps 3;
Matches 293; Conservative 2; Mismatches 6;
QY 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAAATARSVALTISAYQEDPAYLRQCL 55

```

|||||
60 AFLSAHLVAQSLFAYLEHRRVAAAARRSLAKGLDPAATARSVALTISAYQEDPAYLRQCL 119
QY
56 ASARALLYPRARLRVLMVVDGNRAEDLYMDMFREVFADDPATYVMDGNTHQWPEPA-A 114
|||||
120 TSARALLYPHTRLRVLMVVDGNRAEDLYMDMFREVFADDPATYVMDGNTHQWPEPAEA 179
QY
115 AGAVGAGAYREVEADPGRLAVEALVTRRCVCVAQRWGKREVMYTAFAKALGSDVDYVQ 174
|||||
180 TGAVGEGAYREVEADPGRLAVEALVTRRCVCVAQRWGKREVMYTAFAKALGSDVDYVQ 239
QY
175 VCDSDTRLDPMALLELVRLDDEDPRVGAVGGDVRLINPLDSWVSFLSLRYWVAFNVERA 234
DB
240 VCDSDTRLDPMALLELVRLDDEDPRVGAVGGDVRLINPLDSWVSFLSLRYWVAFNVERA 299
QY
235 QSYFHCVCISGPGLYRNLLQOFLAEMYNQKFLGTHCTFGDDRHLTNRMLSMGYATK 294
DB
300 QSYFHCVCISGPGLYRNLLQOFLAEMYNQKFLGTHCTFGDDRHLTNRMLSMGYATK 359
QY
295 -----AEGTRWS 301
DB
360 YTSRRCYSETPSSFLRWLSQOTRWS 385

RESULT 5
ID Q6S742 PRELIMINARY; PRT; 583 AA.
AC Q6S742;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hyaluronan synthase.
GN Name=HAS1;
OS Papio anubis (Olive baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Papio.
OX NCBI_TaxID=9555;
RN [1]
RP SEQUENCE FROM N.A.
RA Martin-Duncker I., Ortol R., Mollicone R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY463695; AAR25554.1; -.
DR InterPro; IPR001173; Glyco_trans_2.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR002057; Isopen_N_synth.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS00185; IPNS_1; UNKNOWN_1.
SQ SEQUENCE 583 AA; 65596 MW; EA47354C89346F94 CRC64;

Query Match 88.9%; Score 1508; DB 2; Length 583;
Best Local Similarity 89.5%; Pred. No. 5.1e-121;
Matches 291; Conservative 4; Mismatches 6; Indels 24; Gaps 2;

QY 1 AFLSAHLVAQSLFAYLEHRRVA-----AARGPLDPAATARSVALTISAYQEDPAYLRQCL 55
DB 61 AFLSAHLVAQSLFAYLEHRRVAAAARRSLAKGLDPAATARSVALTISAYQEDPAYLRQCL 120
QY
56 ASARALLYPRARLRVLMVVDGNRAEDLYMDMFREVFADDPATYVMDGNTHQWPEPAEA 115
DB 121 VSARALLYPRARLRVLMVVDGNRPEDLYMDMFREVFADDPATYVMDGNTHQWPEPAAV 180
QY
116 GAVGAGAYREVEADPGRLAVEALVTRRCVCVAQRWGKREVMYTAFAKALGSDVDYVQV 175
DB 181 GAVGAGAYREVEADPGRLAVEALVTRRCVCVAQRWGKREVMYTAFAKALGSDVDYVQV 240
QY
176 CDSSTRLDPMALLELVRLDDEDPRVGAVGGDVRLINPLDSWVSFLSLRYWVAFNVERAC 235
DB 241 CDSSTRLDPMALLELVRLDDEDPRVGAVGGDVRLINPLDSWVSFLSLRYWVAFNVERAC 300
QY
236 QSYFHCVCISGPGLYRNLLQOFLAEMYNQKFLGTHCTFGDDRHLTNRMLSMGYATK- 294
DB 301 QSYFHCVCISGPGLYRNLLQOFLAEMYNQKFLGTHCTFGDDRHLTNRMLSMGYATKY 360

```

```

|||||
295 -----AEGTRWS 301
DB
361 TSSRCYSETPSSFLRWLSQOTRWS 385

RESULT 6
ID Q8CH93 PRELIMINARY; PRT; 583 AA.
AC Q8CH93;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hyaluronan synthase 1.
GN Name=HAS1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA PubMed=14724275; DOI=10.1074/jbc.M313178200;
RX Itano N., Sawai T., Atsumi F., Miyaishi O., Taniguchi S., Kannagi R.,
RA Hamaguchi M., Kimata K.;
RT "Selective expression and functional characteristics of three
RT Mammalian hyaluronan synthases in oncogenic malignant
RT transformation.";
RL J. Biol. Chem. 279:18679-18687(2004).
DR EMBL; AB097568; BAC43730.1; -.
DR InterPro; IPR001173; Glyco_trans_2.
DR InterPro; IPR002057; Isopen_N_synth.
DR PROSITE; PS00185; IPNS_1; UNKNOWN_1.
SQ SEQUENCE 583 AA; 65724 MW; 35513C6B21DE4E8D CRC64;

Query Match 88.7%; Score 1505.5; DB 2; Length 583;
Best Local Similarity 89.6%; Pred. No. 8.4e-121;
Matches 292; Conservative 2; Mismatches 7; Indels 25; Gaps 3;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDPAATARSVALTISAYQEDPAYLRQCL 55
DB 60 AFLSAHLVAQSLFAYLEHRRVTVAAARRAFAGPLDPAATARSVALTISAYQEDPAYLRQCL 119
QY
56 ASARALLYPRARLRVLMVVDGNRAEDLYMDMFREVFADDPATYVMDGNTHQWPEPA-A 114
DB 120 TSARALLYPRTRLRVLMVVDGNRAEDLYMDMFREVFADDPATYVMDGNTHQWPEPAEA 179
QY
115 AGAVGAGAYREVEADPGRLAVEALVTRRCVCVAQRWGKREVMYTAFAKALGSDVDYVQ 174
DB 180 AGAVGAGAYREVEADPGRLAVEALVTRRCVCVAQRWGKREVMYTAFAKALGSDVDYVQ 239
QY
175 VCDSDTRLDPMALLELVRLDDEDPRVGAVGGDVRLINPLDSWVSFLSLRYWVAFNVERA 234
DB 240 VCDSDTRLDPMALLELVRLDDEDPRVGAVGGDVRLINPLDSWVSFLSLRYWVAFNVERA 299
QY
235 QSYFHCVCISGPGLYRNLLQOFLAEMYNQKFLGTHCTFGDDRHLTNRMLSMGYATK 294
DB 300 QSYFHCVCISGPGLYRNLLQOFLAEMYNQKFLGTHCTFGDDRHLTNRMLSMGYATK 359
QY
295 -----AEGTRWS 301
DB
360 YTSRRCYSETPSSFLRWLSQOTRWS 385

RESULT 7
ID Q6T488 PRELIMINARY; PRT; 458 AA.
AC Q6T488;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE Hyaluronan synthase 1 (Fragment).
GN Name=has1;
OS Brachydanio rerio (Zebrafish) (Danio rerio).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=14729574;
 RA Bakkas J., Kramer C., Pothof J., Quaedvlieg N.E., Spink H.P.,
 RA Hamerschmidt M.;
 RT "Has2 is required upstream of Rac1 to govern dorsal migration of
 RT lateral cells during zebrafish gastrulation.";
 RL Development 131:525-537(2004).
 DR EMBL; AY437407; AAR97372.1; -.
 DR ZFIN; ZDB-GENE-040218-2; has1.
 DR InterPro; IPR001173; Glyco_trans_2.
 FT NON_TER 458 458
 SQ SEQUENCE 458 AA; 53188 MW; 4B421299ABBF96E CRC64;

Query Match 61.8%; Score 1048.5; DB 2; Length 458;
 Best Local Similarity 62.3%; Pred. No. 1.3e-81;
 Matches 195; Conservative 43; Mismatches 56; Indels 19; Gaps 4;

QY 3 LSAHLVAQSLPAYLEHRRVAAARGPLDAATARSVALTISAYOEDPAYLROCLASARALL 62
 DB 57 LSLHVLIOQSFPAFVHGRNARRK---PCSYTKTIGFTISAYOEDPAYLRECLQSVRALQ 113
 QY 63 YPRARLRVLMVVDGNRAEDLYMVDMFREVFADEDPATYVWDGNYPWPAPAA---AGAVG 119
 DB 114 YPSELLRIVMVDGNSEDDRYMLENFRVFAVDQDGCYIWNQNYHS-WNPNGODGGAEG 172
 QY 120 AGAYREBAEDPGRLAVALRTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVCDSD 179
 DB 173 PDADYEVLPFQRLVEVEIIRTKVCIMQKWKGGKREVMYTAFAKALGSSADFIQVCDSD 232
 QY 180 TRLDPMALLRVLRLDEPRVCAVGDVRIINPLDSWTSFLSSLRVWAFNVERACQSYF 239
 DB 233 TKLDPLATVELCKVLESQKYGAVGDMVILNKDSYISFMSLSRYWMAFNVERSCQSPF 292
 QY 240 HCVSCISGPLGLYRNLLQQFLAEMYNQKFLGTHCTFGDDRLHTRNRLSMGYATK---- 294
 DB 293 DCVSCISGPLGLYRNLLQQFLAEMYNQKFLGTHCTFGDDRLHTRNRLSMGYATK 352
 QY 295 -----AEGTRW 300
 DB 353 KCYETPAQFLRW 365

RESULT 8
 HAS1_XENLA
 ID HAS1_XENLA
 AC P13563;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Hyaluronan synthase 1 (EC 2.4.1.212) (Hyaluronate synthase 1)
 DE (Hyaluronic acid synthase 1) (HA synthase 1) (XHASI) (DG42 protein).
 GN Name=HAS1;
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88313363; PubMed=3410156;
 RA Rosa F., Sargent T.D., Rebert M.L., Michaels G.S., Jamrich M.,
 RA Grunz H., Jonas E., Winkles J.A., Dawid I.B.;
 RT "Accumulation and decay of DG42 gene products follow a gradient
 RT pattern during Xenopus embryogenesis.";
 RL Dev. Biol. 129:114-123 (1988).
 CC -!- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA)
 CC synthesis (By similarity). May play a role in signaling or pattern
 CC formation in embryonic development.

CC -!- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-
 CC glucuronate = [beta-N-acetyl-D-glucosaminyl(1->4)beta-D-
 CC glucuronosyl(1->3)](n) + 2n UDP.
 CC -!- COFACTOR: Magnesium.
 CC -!- PATHWAY: Hyaluronate synthesis.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -!- TISSUE SPECIFICITY: Embryo.
 CC -!- DEVELOPMENTAL STAGE: During gastrula and early neurula stages.
 CC -!- SIMILARITY: Belongs to the ncdC/HAS family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M22249; AAA49699.1; -.
 DR PIR; A43740; A43740.
 DR InterPro; IPR001173; Glyco_trans_2.
 KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
 FT DOMAIN 1 28 Cytoplasmic (Potential).
 FT TRANSMEM 29 49 1 (Potential).
 FT DOMAIN 50 61 Extracellular (Potential).
 FT TRANSMEM 62 82 2 (Potential).
 FT DOMAIN 83 411 Cytoplasmic (Potential).
 FT TRANSMEM 412 432 3 (Potential).
 FT DOMAIN 433 433 Extracellular (Potential).
 FT TRANSMEM 434 454 4 (Potential).
 FT DOMAIN 455 456 Cytoplasmic (Potential).
 FT TRANSMEM 457 477 5 (Potential).
 FT DOMAIN 478 505 Extracellular (Potential).
 FT TRANSMEM 506 526 6 (Potential).
 FT DOMAIN 527 543 Cytoplasmic (Potential).
 FT TRANSMEM 544 568 7 (Potential).
 FT DOMAIN 565 588 Extracellular (Potential).
 SQ SEQUENCE 588 AA; 68522 MW; 33DA3B8E331F4CE9 CRC64;

Query Match 56.9%; Score 965.5; DB 1; Length 588;
 Best Local Similarity 60.5%; Pred. No. 2.4e-74;
 Matches 184; Conservative 47; Mismatches 62; Indels 11; Gaps 4;

QY 6 HLVAQSLPAYLEHRRVAAARGPLDAATARSVALTISAYOEDPAYLROCLASARALLYPR 65
 DB 72 HLMWQSLFAFLRIRV---NKSFLPCSKFKTVALTIAGYQENPEYLKCLBCKVYKPK 128
 QY 66 ARLRVLMVVDGNRAEDLYMVDMFREVFADEDPATYVWDGNTHQWPWEPAAAGAVGAYRE 125
 DB 129 DKLKILVIDGNTDDAYMMEMFKDVFHGEDVGTVMKGNHYTVKKPE---ETNKGSCPE 185
 QY 126 VEA---EDPGRLAVALRTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDRL 182
 DB 186 VSKPLNEDEGINMVEELVRNKRKVCIMQWGGKREVMYTAFAIGTSVDYVQVCDSDTKL 245
 QY 183 DPMALLRVLRLDEPRVCAVGDVRIINPLDSWTSFLSSLRVWAFNVERACQSYFHCV 242
 DB 246 DELATVENVKVLESNDMYGAVGDVRIINPLPYDSFISFMSLSRYWMAFNVERACQSYFDCV 305
 QY 243 SCISGPLGLYRNLLQQFLAEMYNQKFLGTHCTFGDDRLHTRNRLSMGYATK--AEGTRW 300
 DB 306 SCISGPLGMRYNNILQVLEAWYRQKFLGTCTGLGDDRLHTRNRLSMGYRTKYTKSRAF 365
 QY 301 SGTP 304
 DB 366 SETP 369

RESULT 9
 Q95M29
 ID Q95M29 PRELIMINARY; PRT; 552 AA.
 AC Q95M29;
 DT 01-DEC-2001 (TRENBLrel. 19, Created)

```
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hyaluronic acid synthase 3.
GN Name=HAS3;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Ohno S.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB05979; BAB63265.1; -.
DR InterPro; IPR001173; Glyco_trans_2.
SQ SEQUENCE 552 AA; 63137 MW; 3C4A805AD5347156 CRC64;

Query Match 55.9%; Score 948.5; DB 2; Length 552;
Best Local Similarity 57.9%; Pred. No. 6.5e-73;
Matches 186; Conservative 39; Mismatches 63; Indels 33; Gaps 4;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDDPAYLRQCLASARA 60
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
50 AILGLHLIIQSLFAPLEHRRRRARRPLKPSRRRSVALCIAAYQEDDPYLRLKCLRSAR 109
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
61 LLYPRARLRVLMVVDGNRAEDLYVDMFREFV-ADDDPATYVWDGNVHQPWEPAAAGAVG 119
Qy :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
110 IAFP--DLKVVVMVDGNQEDAYMLDIFHEVLGCGNQAGFFVWRSNPHEAGEGETEASLQ 167
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
120 AGAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMTAFKALGDSVDYVQVCDSD 179
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
168 EGMR-----VRVVRTSTFCIMQKGGKREVMTAFKALGDSVDYVQVCDSD 216
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
180 TRLDPMALLELRVLDDEPRVAGVGDVRIINPLDSWVSFLSSLRVWAFNVERACQSYF 239
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
217 TVLPDPACTIEMLRVLEEDPQGVGGDVQILNKYDSWISFLSSVRYWMAFNVERACQSYF 276
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
240 HCVSCISGPLGLYRNLLQOFLAAYWYNOKFLGTHCTFGDDRHLTNRMLSMGYATK---- 294
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
277 GCVCISGPLGMYRNLSLQOFLAAYWYNOKFLGSKCSFGDDRHLTNRVLSLGYRTKYTARS 336
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

Qy 295 -----AEGTRWS 301
Db :||||
337 SKCLTETPTKYLRWLNQTRWS 357

RESULT 10
Q75R37 PRELIMINARY; PRT; 553 AA.
AC Q75R37;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Hyaluronic acid synthase 3.
GN Name=has3;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Kimura N., Konno Y., Yokoo M., Sato E.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB159675; BAD18881.1; -.
DR InterPro; IPR001173; Glyco_trans_2.
SQ SEQUENCE 553 AA; 63110 MW; C97EBE911723B44C CRC64;

Query Match 55.3%; Score 938; DB 2; Length 553;
Best Local Similarity 57.5%; Pred. No. 5.2e-72;
Matches 185; Conservative 39; Mismatches 64; Indels 34; Gaps 5;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARG-PLDAATARSVALTISAYQEDDPAYLRQCLASAR 59
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
50 AILGLHLIIQSLFAPLEHRRRRARRPLKPSLQRSVALCIAAYQEDDPYLRLKCLRSAR 109

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDDPAYLRQCLASARA 60
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
51 AFLASHLIIQSLFAPLEHRRKMKKSLETPI--KLNTVALCIAAYQEDDPYLRLKCLQSVKR 108
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
61 LLYPRARLRVLMVVDGNRAEDLYVDMFREFVADDDPATYVWDGNVHQPWEPAAAGAVGA 120
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
109 LTPV--GIKVVWVIDGNSDDLYMMDIFSEVMGRDNTATYWKNNFHEK-----GP 157
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
121 GAYREVEAEDPGRLAVALVTRRCVCVAQRWGGKREVMTAFKALGDSVDYVQVCDSDT 180
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
158 GETDESHKSSQH--VTQLVLSNKSICIMQKGGKREVMTAFRALGRSDYVQVCDSDT 215
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
181 RLDPMALLELRVLDDEPRVAGVGDVRIINPLDSWVSFLSSLRVWAFNVERACQSYFH 240
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
216 MLDPPASSVEMVKVLEEDPQGVGGDVQILNKYDSWISFLSSVRYWMAFNVERACQSYFG 275
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
241 CVSCISGPLGLYRNLLQOFLAAYWYNOKFLGTHCTFGDDRHLTNRMLSMGYATK----- 294
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
276 CVQCISGPLGMYRNLSLHFEVDWYNQEFMGSCQSGFGDDRHLTNRVLSLGYATKYTARSK 335
Qy |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
295 -----AEGTRWS 301
Qy :||||
336 CLTETPTKYLRWLNQTRWS 355
```


Db 109 AQRAPP--NLKVVVVVGNRQEDTYMLDI FHEVLGGTEQAGCFVWRSNFHBAGEGETEA 166
 Qy 117 AVGAGAYREVEAEDPGRLAVALVRTRRCVCVQAQWGGKREVMYTAFAKALGDSVDYVQC 176
 Db 167 SLQEMER-----VRAVWASTFSCIMQKWGGKREVMYTAFAKALGNSVDYIQVC 215
 Qy 177 DSDTRFLDPMALLEVRVLDEDPRGVAGCGDVRILNPLDSWSVFLSLRLVWVAFNVERACQ 236
 Db 216 DSDTVLDPACTIEMLRLEEDPQVGGVGDYQLNKYDSWISFLSVRYMMAFNVERACQ 275
 Qy 237 SYFHCVCISCPGLGYNRLNQFLLEAWYNKFLGTHCTFGDDRHLTNRLMSNGYATK-- 294
 Db 276 SYFCVCVCISCPGLGYNRLNQFLLEAWYNKFLGTHCTFGDDRHLTNRLMSNGYATK 335
 Qy 295 -----AEGTRWS 301
 Db 336 ARSKCLTETPTRYLRWLNLNQTRWS 359

RESULT 13
 HAS2 MOUSE
 ID HAS2 MOUSE STANDARD; PRT; 552 AA.
 AC P70312; P70411; Q62405;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUN-2004 (Rel. 44, Last annotation update)
 DE Hyaluronan synthase 2 (EC 2.4.1.212) (Hyaluronate synthase 2)
 DE (Hyaluronic acid synthase 2) (HA synthase 2).
 GN Name=Has2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=96394433; PubMed=8798545; DOI=10.1074/jbc.271.38.23400;
 RA Spicer A.P., Augustine M.L., McDonald J.A.;
 RT "Molecular cloning and characterization of a putative mouse hyaluronan
 RT synthase.";
 RL J. Biol. Chem. 271:23400-23406 (1996).
 RN [2]
 RP REVISION TO 138.
 RA Spicer A.P., Augustine M.L., McDonald J.A.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBAJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Swiss CD1;
 RX MEDLINE=97169014; PubMed=9016821; DOI=10.1006/abbi.1996.9793;
 RA Fuelep C., Salustri A., Hascall V.C.;
 RT "Coding sequence of a hyaluronan synthase homologue expressed during
 RT expansion of the mouse cumulus-oocyte complex.";
 RL Arch. Biochem. Biophys. 337:261-266 (1997).
 RN [4]
 RP SEQUENCE OF 215-348 FROM N.A.
 RC STRAIN=Swiss Webster; TISSUE=Embryo;
 RX MEDLINE=96209769; PubMed=8643441; DOI=10.1073/pnas.93.10.4548;
 RA Samino C.E., Specht C.A., Raimondi A., Robbins P.W.;
 RT "Homologs of the Xenopus developmental gene DG42 are present in
 RT zebrafish and mouse and are involved in the synthesis of Nod-like
 RT chitin oligosaccharides during early embryogenesis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:4548-4553 (1996).
 CC -1- FUNCTION: Plays a role in hyaluronan/hyaluronic acid (HA)
 CC synthesis.
 CC -1- CATALYTIC ACTIVITY: N UDP-N-acetyl-D-glucosamine + n UDP-D-
 CC glucuronate = [beta-N-acetyl-D-glucosaminyl(1->3)]beta-D-
 CC glucuronosyl(1->3)](n) + 2n UDP.
 CC -1- COFACTOR: Magnesium.
 CC -1- PATHWAY: Hyaluronate synthesis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- TISSUE SPECIFICITY: Expressed in heart, brain, spleen, lung and
 CC skeletal muscle.
 CC -1- DEVELOPMENTAL STAGE: Expressed at day 7.5.

```
CC -!- SIMILARITY: Belongs to the nodC/HAS family.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U52524; AAC53309.2; -.
CC DR EMBL; U69695; AAB17609.1; -.
CC DR EMBL; U53222; AAC52651.1; -.
CC DR MGI; 107821; Haeg2.
CC DR InterPro; IPR001173; Glyco trans 2.
CC DR Pfam; PF00535; Glycos transf 2; 1.
CC KW Glycosyltransferase; Multigene family; Transferase; Transmembrane.
CC FT DOMAIN 1 11
CC FT TRANSMEM 12 32
CC FT DOMAIN 33 45
CC FT TRANSMEM 46 66
CC FT DOMAIN 67 374
CC FT TRANSMEM 375 395
CC FT DOMAIN 396 402
CC FT TRANSMEM 403 423
CC FT DOMAIN 424 429
CC FT TRANSMEM 430 450
CC FT DOMAIN 451 475
CC FT TRANSMEM 476 496
CC FT DOMAIN 497 510
CC FT TRANSMEM 511 531
CC FT DOMAIN 532 552
CC FT CONFLICT 138 138 I -> M (in Ref. 3).
CC SQ SEQUENCE 552 AA; 63492 MW; 7C88019F680DD982 CRC64;

Query Match 55.1%; Score 935; DB 1; Length 552;
Best Local Similarity 55.0%; Pred. No. 9.4e-72;
Matches 176; Conservative 50; Mismatches 60; Indels 34; Gaps 5;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYOEDPAYLRQCLASARA 60
Db 51 AFLASHLLIQSLFAPLEHRRKMKKSLETPI--KLNKTVALCIAAYQEDPDYLRKCLQSVKR 108

QY 61 LLYPRARLRVLVVDGNRAEDLYVDMFREVFADEDPATYVWDGNYHQPWPAPAAAGVGA 120
Db 109 LTYP--GIKVVWVIDGNSDDLYMMDIFSEVIGRDKSATYIWKNNFHEK-----GP 157

QY 121 GAYREVEAEDPGRLAVALVTRRCVCAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 180
Db 158 GETESHKESQ--VTQLVLSNKSICIMQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 215

QY 181 RLDPMALLELRVLDEDPVGAAGDVRILNPLDSWVSFLSSLYWVAFNVERACQSYFH 240
Db 216 MLDPASSVEMVKVLEEDPMVGVDVQILNKYDSWISFLSSVRYWVAFNVERACQSYFG 275

QY 241 CVCSISGPLGLYRNLLQOFLFVWYKFLGTHCTFGDDRLHTRNRLSMGYATK----- 294
Db 276 CVQCISGPLGLYRNLLHFEVWYKFLGTHCTFGDDRLHTRNRLSLGATKYTARSK 335

QY 295 -----ABGTRWS 301
Db 336 CLTETPIEYLRWLNLQOTRWS 355

RESULT 15
ID Q95M16 PRELIMINARY; PRT; 552 AA.
AC Q95M16;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hyaluronic acid synthase 2.
GN Name=HAS2;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21363038; PubMed=11470161; DOI=10.1016/S0167-4781(01)00256-1;
RA Ohno S., Tanimoto K., Fujimoto K., Ijuin C., Honda K., Tanaka N.,
RA Doi T., Nakahara M., Tanne K.;
RT "Molecular cloning of rabbit hyaluronic acid synthases and their
RT expression patterns in synovial membrane and articular cartilage.";
RL Biochim. Biophys. Acta 1520:71-78(2001).
DR EMBL; AB055978; BAB63264.1; -.
DR InterPro; IPR001173; Glyco trans 2.
DR SQ SEQUENCE 552 AA; 63483 MW; 4A1A6E6960B0B7B1 CRC64;

Query Match 55.1%; Score 935; DB 2; Length 552;
Best Local Similarity 55.3%; Pred. No. 9.4e-72;
Matches 177; Conservative 49; Mismatches 60; Indels 34; Gaps 5;

QY 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYOEDPAYLRQCLASARA 60
Db 51 AFLASHLLIQSLFAPLEHRRKMKKSLETPI--KLNKTVALCIAAYQEDPDYLRKCLQSVKR 108

QY 61 LLYPRARLRVLVVDGNRAEDLYVDMFREVFADEDPATYVWDGNYHQPWPAPAAAGVGA 120
Db 109 LTYP--GIKVVWVIDGNSDDLYMMDIFSEVIGRDKSATYIWKNNFHEK-----GP 157

QY 121 GAYREVEAEDPGRLAVALVTRRCVCAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 180
Db 158 GETESHKESQ--VTQLVLSNKSICIMQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 215

QY 181 RLDPMALLELRVLDEDPVGAAGDVRILNPLDSWVSFLSSLYWVAFNVERACQSYFH 240
Db 216 MLDPASSVEMVKVLEEDPMVGVDVQILNKYDSWISFLSSVRYWVAFNVERACQSYFG 275

QY 241 CVCSISGPLGLYRNLLQOFLFVWYKFLGTHCTFGDDRLHTRNRLSMGYATK----- 294
Db 276 CVQCISGPLGLYRNLLHFEVWYKFLGTHCTFGDDRLHTRNRLSLGATKYTARSK 335

QY 295 -----ABGTRWS 301
Db 336 CLTETPIEYLRWLNLQOTRWS 355

RESULT 14
ID Q8HZJ3 PRELIMINARY; PRT; 552 AA.
AC Q8HZJ3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hyaluronic synthase 2.
OS Equus caballus (Horse).
```

```
QY 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 51 AFLASHLIQSLFAFLHRRKMKKSLTPI--KLNTVALCIAAYQEDPDYLRKCLQSVKR 108
QY 61 LLYPEARLRLVLMVDGNRAEDLYMVMREVPADDPATYVWDGNHYQHPWEPAAGAVGA 120
Db 109 LLYP--GIKVMVINDGNSEDDVIMMDIFSEVNGRETSATYIWKNNFHEK-----GP 157
QY 121 GAYREVEAEDPGRLAVEALVTRRCVCVAQRWGGKREVMYTAFAKALGDSVDYVQVCDSDT 180
Db 158 GETDESHKESQH--VTQLVLSNKSVCIMQKWGGKREVMYTAFAKALGDSVDYVQVCDSDT 215
QY 181 RLDPWALLELVRLDEDPVGVAGGVDVRIILNPLDSWVSLSLRYWVAFNVERACQSYFH 240
Db 216 MLDPASSVEMVXVLEEDPMVGGVGGVQILNKYDSWISFLSSVRYWMAFNTERACQSYFG 275
QY 241 CVSCISGPIGLYRNLLQOFLAAYNQKPLGTHCTFGDDRHLTNRLSMGYATK----- 294
Db 276 CVQCISGPIGMYRNSLLHEFVEDWYNQBFMGNQCSFGDDRHLTNRVLSLGYATKYTARSK 335
QY 295 -----AEGTRWS 301
Db 336 CLTETPIEYLRWLNQOTRWS 355
```

Search completed: March 11, 2005, 14:22:04
Job time : 89.6364 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 14:11:23 ; Search time 23.3333 Seconds
(without alignments)
1023.759 Million cell updates/sec

Title: US-10-672-399-8

Perfect score: 1697

Sequence: 1 AFLSAHLVAOSLFAYLEHRR.....SGTPGPAAYTORPRPSCGG 320

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/1aa/5A COMB.pep.*
- 2: /cgn2_6/ptodata/1/1aa/5B COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/6A COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6B COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/PTUS COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1553.5	91.5	578	US-08-635-552A-2	Sequence 2, Appli
2	1553.5	91.5	589	US-09-949-016-11530	Sequence 11530, A
3	1550.5	91.4	543	US-09-155-768-4	Sequence 4, Appli
4	1508.5	88.9	583	US-08-675-499A-3	Sequence 3, Appli
5	1508.5	88.9	583	US-08-812-008-3	Sequence 3, Appli
6	965.5	56.9	587	US-08-635-552A-3	Sequence 3, Appli
7	949	55.9	587	US-08-675-499A-4	Sequence 4, Appli
8	949	55.9	587	US-08-812-008-4	Sequence 4, Appli
9	935.5	55.1	554	US-08-812-008-32	Sequence 32, Appli
10	935	55.1	552	US-08-675-499A-2	Sequence 2, Appli
11	935	55.1	552	US-08-812-008-2	Sequence 2, Appli
12	934	55.0	552	US-09-949-016-6608	Sequence 6608, Ap
13	928	54.7	553	US-09-949-016-9599	Sequence 9599, Ap
14	415.5	24.5	241	US-08-865-173-2	Sequence 2, Appli
15	415.5	24.5	241	US-09-385-174-2	Sequence 2, Appli
16	409	24.1	568	US-09-469-200E-10	Sequence 10, Appli
17	279	16.4	426	US-08-675-499A-6	Sequence 6, Appli
18	279	16.4	426	US-08-812-008-6	Sequence 6, Appli
19	268.5	15.8	190	US-08-812-008-29	Sequence 29, Appli
20	268.5	15.8	190	US-08-812-008-30	Sequence 30, Appli
21	265	15.6	78	US-08-675-499A-24	Sequence 24, Appli
22	265	15.6	78	US-08-675-499A-27	Sequence 27, Appli
23	265	15.6	78	US-08-675-499A-28	Sequence 28, Appli
24	265	15.6	78	US-08-812-008-24	Sequence 24, Appli
25	265	15.6	78	US-08-812-008-27	Sequence 27, Appli
26	265	15.6	78	US-08-812-008-28	Sequence 28, Appli
27	231.5	13.6	417	US-09-469-200E-2	Sequence 2, Appli

28	217	12.8	43	4	US-08-675-499A-9	Sequence 9, Appli
29	217	12.8	43	4	US-08-812-008-9	Sequence 9, Appli
30	213.5	12.6	395	4	US-08-635-552A-4	Sequence 4, Appli
31	213.5	12.6	419	2	US-08-270-581-2	Sequence 2, Appli
32	213.5	12.6	419	4	US-09-146-893-2	Sequence 2, Appli
33	213.5	12.6	419	4	US-08-675-499A-5	Sequence 5, Appli
34	213.5	12.6	419	4	US-08-812-008-5	Sequence 5, Appli
35	177	10.4	43	4	US-08-812-008-35	Sequence 35, Appli
36	173	10.2	43	4	US-08-675-499A-7	Sequence 7, Appli
37	173	10.2	43	4	US-08-812-008-7	Sequence 7, Appli
38	163	9.6	43	4	US-08-675-499A-10	Sequence 10, Appli
39	163	9.6	43	4	US-08-812-008-10	Sequence 10, Appli
40	160	9.4	403	4	US-09-902-540-11529	Sequence 11529, A
41	153.5	9.0	1093	4	US-09-248-796A-17108	Sequence 17108, A
42	143.5	8.5	416	2	US-08-867-030B-12	Sequence 12, Appli
43	143.5	8.5	416	5	PCT-US95-06119-12	Sequence 12, Appli
44	133.5	7.9	393	4	US-09-902-540-11514	Sequence 11514, A
45	130.5	7.7	55	4	US-08-675-499A-29	Sequence 29, Appli

ALIGNMENTS

RESULT 1
US-08-635-552A-2
; Sequence 2, Application US/08635552A
; Patent No. 6423514
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; TITLE OF INVENTION: Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS98-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 578 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-635-552A-2

Query Match	91.5%;	Score 1553.5;	DB 4;	Length 578;
Best Local Similarity	93.1%;	Pred. No. 4.9e-173;		
Matches 298;	Conservative 2;	Mismatches 1;	Indels 19;	Gaps 1;
Oy	1	AFLSAHLVAOSLFAYLEHRRVAAAARGPLOAATARSVALTISAYQEPAYLRQCLASARA	60	
Db	61	AFLSAHLVAOSLFAYLEHRRVAAAARGPLOAATARSVALTISAYQEPAYLRQCLASARA	120	
Oy	61	LLYPRLRLVLMVDGNRAEDLYNVDMPREVFADDEPATYVWDGNYHQHPWEPAAGAYGA	120	

Db 121 LLYPRARLRVLMVVDGNRAEDLYMDFREVFADDPATYVWDGNYHQWEPAAAAAGVGA 180
Qy 121 GAYREVEADPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 180
Db 181 GAYREVEADPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 240
Qy 181 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWSFSLSSRYWAFNVERACQSYFH 240
Db 241 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWSFSLSSRYWAFNVERACQSYFH 300
Qy 241 CVSCISGPLGLYRNLLQQFLFLEAWYNQKFLGTHCTFGDDRHLTNRMLSMGVATK----- 294
Db 301 CVSCISGPLGLYRNLLQQFLFLEAWYNQKFLGTHCTFGDDRHLTNRMLSMGVATKTSRSR 360
Qy 295 -----AEGTRWS 301
Db 361 CYSETPSSFLRWLSQQTWS 380

RESULT 2

US-09-949-016-11530
; Sequence 11530, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASES, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11530
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11530

Query Match 91.5%; Score 1553.5; DB 4; Length 589;
Best Local Similarity 93.1%; Pred. No. 5.1e-173;
Matches 298; Conservative 2; Mismatches 1; Indels 19; Gaps 1;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 72 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 131
Qy 61 LLYPRARLRVLMVVDGNRAEDLYMDFREVFADDPATYVWDGNYHQWEPAAAAAGVGA 120
Db 132 LLYPRARLRVLMVVDGNRAEDLYMDFREVFADDPATYVWDGNYHQWEPAAAAAGVGA 191
Qy 121 GAYREVEADPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 180
Db 192 GAYREVEADPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 251
Qy 181 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWSFSLSSRYWAFNVERACQSYFH 240
Db 252 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWSFSLSSRYWAFNVERACQSYFH 311
Qy 241 CVSCISGPLGLYRNLLQQFLFLEAWYNQKFLGTHCTFGDDRHLTNRMLSMGVATK----- 294
Db 312 CVSCISGPLGLYRNLLQQFLFLEAWYNQKFLGTHCTFGDDRHLTNRMLSMGVATKTSRSR 371
Qy 295 -----AEGTRWS 301
Db 372 CYSETPSSFLRWLSQQTWS 391

RESULT 3

US-09-155-768-4
; Sequence 4, Application US/09155768A
; Patent No. 6162508
; GENERAL INFORMATION:
; APPLICANT: SEIKAGAKU CORPORATION
; TITLE OF INVENTION: POLYPEPTIDE OF HYALURONAN SYNTHASE OF
HUMAN-ORIGIN AND DNA ENCODING THE SAME
; FILE REFERENCE: TOYAM37,001APC
; CURRENT APPLICATION NUMBER: US/09/155,768A
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: JP 8-084326
; EARLIER FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: JP 8-109663
; EARLIER FILING DATE: 1996-04-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 543
; TYPE: PRT
; ORGANISM: HUMAN
US-09-155-768-4

Query Match 91.4%; Score 1550.5; DB 3; Length 543;
Best Local Similarity 92.8%; Pred. No. 1e-172;
Matches 297; Conservative 3; Mismatches 1; Indels 19; Gaps 1;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 26 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 85
Qy 61 LLYPRARLRVLMVVDGNRAEDLYMDFREVFADDPATYVWDGNYHQWEPAAAAAGVGA 120
Db 86 LLYPRARLRVLMVVDGNRAEDLYMDFREVFADDPATYVWDGNYHQWEPAAAAAGVGA 145
Qy 121 GAYREVEADPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 180
Db 146 GAYREVEADPGRLAVALVTRRCVCVAQWGGKREVMYTAFAKALGDSVDYVQVCDSDT 205
Qy 181 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWSFSLSSRYWAFNVERACQSYFH 240
Db 206 RLDPMALLELRVRLDEDPVGAAGDVRILNPLDSWSFSLSSRYWAFNVERACQSYFH 265
Qy 241 CVSCISGPLGLYRNLLQQFLFLEAWYNQKFLGTHCTFGDDRHLTNRMLSMGVATK----- 294
Db 266 CVSCISGPLGLYRNLLQQFLFLEAWYNQKFLGTHCTFGDDRHLTNRMLSMGVATKTSRSR 325
Qy 295 -----AEGTRWS 301
Db 326 CYSETPSSFLRWLSQQTWS 345

RESULT 4

US-08-675-499A-3
; Sequence 3, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P. O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/675,499A
FILING DATE: 03-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 150.170US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-359-3260
TELEFAX: 612-359-3263
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-675-499A-3

Query Match 88.9%; Score 1508.5; DB 4; Length 583;
Best Local Similarity 89.9%; Pred. No. 9.5e-168;
Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;
QY 1 AFLSHLVQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
DB 60 AFLSHLVQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
QY 56 ASARALLYPRARLRVLMVVDGNRAEDLYVMDMREVFADDPATYVWDGNHYQHPWEPA-A 114
DB 120 TSARALLYPHTRLRVLMVVDGNRAEDLYVMDMREVFADDPATYVWDGNHYQHPWEPA 179
QY 115 AGAVGAGAYREVEADPGRLAVEALVTRRCVCVQAQRWGGKREVVMYTAFAKALGDSVDYVQ 174
DB 180 TGAVEGAYREVEADPGRLAVEALVTRRCVCVQAQRWGGKREVVMYTAFAKALGDSVDYVQ 239
QY 175 VCDSDTRLDPMALLELRVLDDEPRVAGVGGDVRILNPLDSWVSLSLRYVWAFNVRA 234
DB 240 VCDSDTRLDPMALLELRVLDDEPRVAGVGGDVRILNPLDSWVSLSLRYVWAFNVRA 299
QY 235 COSYFHCVCISGPIGLYRNLLQOFLEAWYNQKFLGTHCTFGDDRHLTNRLMSGYATK 294
DB 300 COSYFHCVCISGPIGLYRNLLQOFLEAWYNQKFLGTHCTFGDDRHLTNRLMSGYATK 359
QY 295 -----AEGTRWS 301
DB 360 YTSRRCYSETPSSFLRWLSQOTRWS 385

RESULT 5
US-08-812-008-3
Sequence 3, Application US/08812008
Patent No. 6602693
GENERAL INFORMATION:
APPLICANT: McDonald, J. A.
APPLICANT: Spicer, A. P.
APPLICANT: Augustine, M. L.
TITLE OF INVENTION: GENE ENCODING HYALURONAN
NUMBER OF INVENTION: SYNTHASE
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: P.O. Box 2938
CITY: Minneapolis
STATE: MN
COUNTRY: U.S.A
ZIP: 55402

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/812,008
FILING DATE: 05-MAR-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/675,499
FILING DATE: 03-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Embretson, Janet E
REGISTRATION NUMBER: 39,665
REFERENCE/DOCKET NUMBER: 150.183US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-373-6900
TELEFAX: 612-339-3061
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-812-008-3
Query Match 88.9%; Score 1508.5; DB 4; Length 583;
Best Local Similarity 89.9%; Pred. No. 9.5e-168;
Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;
QY 1 AFLSHLVQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
DB 60 AFLSHLVQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
QY 56 ASARALLYPRARLRVLMVVDGNRAEDLYVMDMREVFADDPATYVWDGNHYQHPWEPA-A 114
DB 120 TSARALLYPHTRLRVLMVVDGNRAEDLYVMDMREVFADDPATYVWDGNHYQHPWEPA 179
QY 115 AGAVGAGAYREVEADPGRLAVEALVTRRCVCVQAQRWGGKREVVMYTAFAKALGDSVDYVQ 174
DB 180 TGAVEGAYREVEADPGRLAVEALVTRRCVCVQAQRWGGKREVVMYTAFAKALGDSVDYVQ 239
QY 175 VCDSDTRLDPMALLELRVLDDEPRVAGVGGDVRILNPLDSWVSLSLRYVWAFNVRA 234
DB 240 VCDSDTRLDPMALLELRVLDDEPRVAGVGGDVRILNPLDSWVSLSLRYVWAFNVRA 299
QY 235 COSYFHCVCISGPIGLYRNLLQOFLEAWYNQKFLGTHCTFGDDRHLTNRLMSGYATK 294
DB 300 COSYFHCVCISGPIGLYRNLLQOFLEAWYNQKFLGTHCTFGDDRHLTNRLMSGYATK 359
QY 295 -----AEGTRWS 301
DB 360 YTSRRCYSETPSSFLRWLSQOTRWS 385

RESULT 6
US-08-635-552A-3
Sequence 3, Application US/08635552A
Patent No. 6423514
GENERAL INFORMATION:
APPLICANT: Briekin, Michael J.
TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA


```
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-635-552A-3

Query Match 56.9%; Score 965.5; DB 4; Length 587;
Best Local Similarity 60.5%; Pred. No. 5.3e-104;
Matches 184; Conservative 47; Mismatches 62; Indels 11; Gaps 4;

Qy 6 HLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPR 65
Db 72 HLMQSLFAFLAIRRV---NKSELPCKFKTVALTIAGYQENPEYLIKLESCKVVKPK 128

Qy 66 ARLRVLMVDGNRAEDLYMDFREVFADDEDPATYVWDGNHQHWPFAAGAGAGAYRE 125
Db 129 DKLKIILVIDGNTDDAYMMEMFKDVFHGEDVGTVMKGNHYTVKKPE---ETNKGSCPE 185

Qy 126 VEA---EDPGLRAVEALVTRRCVCVAQRWGKREVMYTAFKALGDSVDYVQVCDSTRL 182
Db 186 VSKPLNEDEGINVVEELVRNRCVCVIMQWQKREVMYTAFAQIGTSVDYVQVCDSTKL 245

Qy 183 DPMALLERVLDEDPVRGAGGVDRILNPLDSWVSFLSLRYWAFNVERACOSYFHCY 242
Db 246 DELATVEMVKVLESNDMTYGAVGVDVRLNPNVDSFISFMSLRYYWAFNVERACOSYFDCV 305

Qy 243 SCISGPLGLYRNLLIQFLEAWYQKFLGTCTHCTFGDDRHLLNRMLSMGYATK--AEGTRW 300
Db 306 SCISGPLGMRYNNILQVLEAWYRQKFLGTCTYCTGLGDDRHLLNRVLSMGYRTKYTHKSRAP 365

Qy 301 SGTP 304
Db 366 SETP 369

RESULT 7
US-08-675-499A-4
; Sequence 4, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwesman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402

; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-675-499A-4

Query Match 55.9%; Score 949; DB 4; Length 587;
Best Local Similarity 60.2%; Pred. No. 4.6e-102;
Matches 183; Conservative 47; Mismatches 62; Indels 12; Gaps 5;

Qy 6 HLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPR 65
Db 72 HLMQSLFAFLAIRRV---NKSELPCKFKTVALTIAGYQENPEYLIKLESCKVVKPK 128

Qy 66 ARLRVLMVDGNRAEDLYMDFREVFADDEDPATYVWDGNHQHWPFAAGAGAGAYRE 125
Db 129 DKLKIILVIDGNTDDAYMMEMFKDVFHGEDVGTVMKGNHYTVKKPE---ETNKGSCPE 185

Qy 126 VEA---EDPGLRAVEALVTRRCVCVAQRWGKREVMYTAFKALGDSVDYVQVCDSTRL 182
Db 186 VSKPLNEDEGINVVEELVRNRCVCVIMQWQKREVMYTAFAQIGTSVDYVQVCDSTKL 244

Qy 183 DPMALLERVLDEDPVRGAGGVDRILNPLDSWVSFLSLRYWAFNVERACOSYFHCY 242
Db 245 DELATVEMVKVLESNDMTYGAVGVDVRLNPNVDSFISFMSLRYYWAFNVERACOSYFDCV 304

Qy 243 SCISGPLGLYRNLLIQFLEAWYQKFLGTCTHCTFGDDRHLLNRMLSMGYATK--AEGTRW 300
Db 305 SCISGPLGMRYNNILQVLEAWYRQKFLGTCTYCTGLGDDRHLLNRVLSMGYRTKYTHKSRAP 364

Qy 301 SGTP 304
Db 365 SETP 368

RESULT 8
US-08-812-008-4
; Sequence 4, Application US/08812008
; Patent No. 6602693
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwesman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
```

```

; STATE: MN
; COUNTRY: U.S.A
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,008
; FILING DATE: 05-MAR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/675,499
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Embretson, Janet E
; REGISTRATION NUMBER: 39,665
; REFERENCE/DOCKET NUMBER: 150.183US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-812-008-4

Query Match 55.9%; Score 949; DB 4; Length 587;
Best Local Similarity 60.2%; Pred. No. 4.6e-102; Indels 12; Gaps 5;
Matches 183; Conservative 47; Mismatches 62;

Qy 6 HLVAQSLPAYLEHRRVAAAARGPLDAAATARSVALTISAYQEDPAYLRQCLASARALLYPR 65
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
72 HLMVQSLPAFLERRV---NKSELPSPKKTVALTIAGYQENPEYLIKLESCKVYKYPK 128
Qy 66 ARLRLVMVVDGNRAEDLYMVDMPREVFADDPATYVWDGNTHQHPWEPAAAGVAGAYRE 125
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
129 DLKILVIDGNTEDDAYMMEMFKDVFGEVDVGTVMKGNVHTVKKPE---ETNKGSCPE 185
Qy 126 VEA---EDPGRLAVEALVTRRCVCAQWGGKREVMYTAFAKALGDSVDYVQVCDSDTRL 182
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
186 VSKPLNEDEGINVWEELVRNRCVCIMQW-GKREVMYTAFAKALGDSVDYVQVCDSDTKL 244
Qy 183 DPMALLELRVLDEDPVGVAGVGDVRIINPLDWSVFLSSRLRYWVAFNVERACOSYFHCV 242
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
245 DELATVENVKVLESNDMYGAVGGDVRIINPYDSFISFMSLSRLRYWMAFNVERACQSYFDCV 304
Qy 243 SCISGPLGLYRNLLQQFLAWYNQKFLGTCTGCGDDRHLLTNRMLSMGYATK--AEGTRW 300
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
305 SCISGPLGMYRNLLQVLEAWYRQKFLGTCTGCGDDRHLLTNRMLSMGYATK--AEGTRW 364
Qy 301 SGTP 304
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
365 SETP 368

RESULT 9
US-08-812-008-32
; Sequence 32, Application US/08812008
; Patent No. 6602693
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustine, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Schwegman, Lundberg, Moesner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: U.S.A
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,008
; FILING DATE: 05-MAR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/675,499
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Embretson, Janet E
; REGISTRATION NUMBER: 39,665
; REFERENCE/DOCKET NUMBER: 150.183US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 554 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-812-008-32

Query Match 55.1%; Score 935.5; DB 4; Length 554;
Best Local Similarity 57.1%; Pred. No. 1.6e-100;
Matches 185; Conservative 41; Mismatches 61; Indels 37; Gaps 6;

Qy 1 AFLSAHLVAQSLPAYLEHRRVAAAARGPLD---AATARSVALTISAYQEDPAYLRQCLAS 57
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
50 AILGLHLLIQLSFLAFLEHRRMRAGR-PLKLHCSQRSVALCIAAYQEDPEYLRKCLRS 108
Qy 58 ARALLYPRARLRVMVVDGNRAEDLYMVDMPREVF-ADDPATYVWDGNTHQHPWEPAAAG 116
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
109 AQRATFP--NLKVVVVDGNHROEDTYMLDIFHEVLGTEQAGFFWRSNFEAGEGETEA 166
Qy 117 AVGAGAYREVEAEDPGRLAVEALVTRRCVCAQWGGKREVMYTAFAKALGDSVDYVQVC 176
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 SLQEGMER-----VRAVVWASTFSCIMQKMGKREVMYTAFAKALGNSVDYIQVC 215
Qy 177 DSDTRLPDPMALLELRVLDEDPVGVAGVGDVRIINPLDWSVFLSSRLRYWVAFNVERACQ 236
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 DSDTVLPDPACTIEMLRVLEEDPQVGGVGDVQILNKYDSDWSFLSSRVNMAFNVERACQ 275
Qy 237 SYFHCVCSISGPLGLYRNLLQQFLAWYNQKFLGTCTGCGDDRHLLTNRMLSMGYATK-- 294
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
276 SYFGVCVCISGPLGMYRNLLQQFLAWYNQKFLGTCTGCGDDRHLLTNRMLSMGYATK-- 335
Qy 295 -----AEGTRWS 301
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
336 ARSKCLTETPTTRYLRWLNLNQTRWS 359

RESULT 10
US-08-675-499A-2
; Sequence 2, Application US/08675499A
; Patent No. 6492150
; GENERAL INFORMATION:
; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustin, M. L.

```

```
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; TITLE OF INVENTION: SYNTHASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESS: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,499A
; FILING DATE: 03-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Woessner, Warren D
; REGISTRATION NUMBER: 30,440
; REFERENCE/DOCKET NUMBER: 150.170US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-359-3260
; TELEFAX: 612-359-3263
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-675-499A-2

Query Match 55.1%; Score 935; DB 4; Length 552;
Best Local Similarity 55.0%; Pred. No. 1.8e-100;
Matches 176; Conservative 50; Mismatches 60; Indels 34; Gaps 5;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 51 AFLASHLIQSLFAPLEHRRKMKKSLETPI--KLKNTVALCIAAYQEDPDYLRKCLQSVKR 108
Qy 61 LLYPRARLRVLMVVDGNRAEDLYMVDMPREVFADEDPATYVWDGNYHQWPEPAAAGAVGA 120
Db 109 LTYP--GIKVMVVDGNSDDLLYMDIFSEVIGRDKSATYWKKNFHEK-----GP 157
Qy 121 GAYREVEAEDPGRLAVALVTRRCVCAQWGGKREVMYTFAPKALGDSVDYVQVCDSDT 180
Db 158 GETEESHKESQH--VTQLVLSNKSICIMQKMGKREVMYTFAPALGRSDYVQVCDSDT 215
Qy 181 RLDPMALLEVRVLDEDPYRGVAGDVRILNPLDSWISFLSSRYWYAFNVERACQSYFH 240
Db 216 MLDPASSVEMVKVLEEDPMVGGVDVQILNKYDSWISFLSSRYWYAFNVERACQSYFG 275
Qy 241 CVCSIGPLGLYRNLLQQLFLEAWYKQKFLGTHCTFGDDRLHTRNRLSMGYATK----- 294
Db 276 CVQCISGPLGMYRNLHFEFVEDWYNOBFMGNCQSGFDDRLHTRNRLSLGYATKYTARSK 335
Qy 295 -----AEGTRWS 301
Db 336 CLTETPIEYLRWLNQOTRWS 355

RESULT 11
US-08-812-008-2
; Sequence 2, Application US/08812008
; Patent No. 6602693
; GENERAL INFORMATION:

; APPLICANT: McDonald, J. A.
; APPLICANT: Spicer, A. P.
; APPLICANT: Augustine, M. L.
; TITLE OF INVENTION: GENE ENCODING HYALURONAN
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
; STREET: P.O. Box 2938
; CITY: Minneapolis
; STATE: MN
; COUNTRY: U.S.A
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/812,008
; FILING DATE: 05-MAR-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/675,499
; FILING DATE: 03-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Embretson, Janet E
; REGISTRATION NUMBER: 39,665
; REFERENCE/DOCKET NUMBER: 150.183US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-373-6900
; TELEFAX: 612-339-3061
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 552 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-812-008-2

Query Match 55.1%; Score 935; DB 4; Length 552;
Best Local Similarity 55.0%; Pred. No. 1.8e-100;
Matches 176; Conservative 50; Mismatches 60; Indels 34; Gaps 5;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 51 AFLASHLIQSLFAPLEHRRKMKKSLETPI--KLKNTVALCIAAYQEDPDYLRKCLQSVKR 108
Qy 61 LLYPRARLRVLMVVDGNRAEDLYMVDMPREVFADEDPATYVWDGNYHQWPEPAAAGAVGA 120
Db 109 LTYP--GIKVMVVDGNSDDLLYMDIFSEVIGRDKSATYWKKNFHEK-----GP 157
Qy 121 GAYREVEAEDPGRLAVALVTRRCVCAQWGGKREVMYTFAPKALGDSVDYVQVCDSDT 180
Db 158 GETEESHKESQH--VTQLVLSNKSICIMQKMGKREVMYTFAPALGRSDYVQVCDSDT 215
Qy 181 RLDPMALLEVRVLDEDPYRGVAGDVRILNPLDSWISFLSSRYWYAFNVERACQSYFH 240
Db 216 MLDPASSVEMVKVLEEDPMVGGVDVQILNKYDSWISFLSSRYWYAFNVERACQSYFG 275
Qy 241 CVCSIGPLGLYRNLLQQLFLEAWYKQKFLGTHCTFGDDRLHTRNRLSMGYATK----- 294
Db 276 CVQCISGPLGMYRNLHFEFVEDWYNOBFMGNCQSGFDDRLHTRNRLSLGYATKYTARSK 335
Qy 295 -----AEGTRWS 301
Db 336 CLTETPIEYLRWLNQOTRWS 355

RESULT 12
US-09-949-016-6608
```

; Sequence 6608, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6608
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6608

Query Match 55.0%; Score 934; DB 4; Length 552;
Best Local Similarity 55.0%; Pred. No. 2.4e-100;
Matches 176; Conservative 50; Mismatches 60; Indels 34; Gaps 5;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARG-PLDAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 51 AFLASHLLIQSLFAPLEHRRKWKSLTFI--KLNTVALCIAAQQEDPDYLRKCLQSVKR 108
Qy 61 LLYPRARLRVLMVVDGNRAEDLYVDMFREFVADDPATYVVDGNYHQWPEPAAAGAVGA 120
Db 109 LLYP--GIKVVWVIDGNEDDLVMYDIFSEVWGRDKSATYIKNNFHEK-----GP 157
Qy 121 GAYREVEAEDPGRLAVALVTRRCVCVQAORGGKREVMYTAFAKALGDSVDYVQVCDSDT 180
Db 158 GETDESHKESQ--VTQLVLSNKSICIMQKGGKREVMYTAFAKALGDSVDYVQVCDSDT 215
Qy 181 RLDPMLLELVRVLDEDPVAGVGDVRLINPLDSWVSFLSLRYVWAFNVERACQSYF 240
Db 216 MLDPASSVEMWVLEEDPMVGGVGDVQILNKYDSWISFLSVRYWMAFNVERACQSYF 275
Qy 241 CVCSISGPGLYRNLLQOFLQWYKPLGTHCTFGDDRLHNLNRLSMGYATK----- 294
Db 276 CVQCISGPGLYRNLLQOFLQWYKPLGTHCTFGDDRLHNLNRLSMGYATK----- 335
Qy 295 -----AEGTRWS 301
Db 336 CLTETPIEYLRWLNQOTRWS 355

RESULT 13
US-09-949-016-9599
; Sequence 9599, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9599
; LENGTH: 553

; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9599

Query Match 54.7%; Score 928; DB 4; Length 553;
Best Local Similarity 56.8%; Pred. No. 1.2e-99;
Matches 183; Conservative 40; Mismatches 65; Indels 34; Gaps 5;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARG-PLDAATARSVALTISAYQEDPAYLRQCLASARA 59
Db 50 AILGLHLLIQSLFAPLEHRRMRAGQALKJPSRGSVALCIAAQEDPDYLRKCLRAQ 109
Qy 60 ALLYPRARLRVLMVVDGNRAEDLYVDMFREFV-ADDPATYVVDGNYHQWPEPAAAGAV 118
Db 110 RISFP--DLKVVWVVDGNRQEDAYMLDIFHEVLGGTEQAGFPVWRSNFHEAGEGETEASL 167
Qy 119 GAGAYREVEAEDPGRLAVALVTRRCVCVQAORGGKREVMYTAFAKALGDSVDYVQVCDSD 178
Db 168 QEGMDR-----VRDVVRASTFCINQKGGKREVMYTAFAKALGDSVDYVQVCDSD 216
Qy 179 DTRLDPMALLVRVLDEDPVAGVGDVRLINPLDSWVSFLSLRYVWAFNVERACQSY 238
Db 217 DTVLDPACTIEMLRVLEEDPMVGGVGDVQILNKYDSWISFLSVRYWMAFNVERACQSY 276
Qy 239 FHCVCISGPGLYRNLLQOFLQWYKPLGTHCTFGDDRLHNLNRLSMGYATK---- 294
Db 277 FGCVCISGPGLYRNLLQOFLQWYKPLGTHCTFGDDRLHNLNRLSMGYATK---- 336
Qy 295 -----AEGTRWS 301
Db 337 SKCLTETPTKYLRWLNQOTRWS 358

RESULT 14
US-08-865-273-2
; Sequence 2, Application US/08865273
; Patent No. 5994100
; GENERAL INFORMATION:
; APPLICANT: ZHU, YUAN
; APPLICANT: NAMBI, PONNAL
; APPLICANT: PULLEN, MARK A
; TITLE OF INVENTION: NOVEL HAS2 SPLICING VARIANT
; TITLE OF INVENTION: HOEFC11: A TARGET IN CHRONIC RENAL FAILURE,
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,273
; FILING DATE: 29-MAY-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:

INFLAMMAT

Result No.	Score	Query			ID	Description
		Match	Length	DB		
1	1697	100.0	320	16	US-10-672-399-8	Sequence 8, Appli
2	1553.5	91.5	578	13	US-10-042-523-2	Sequence 2, Appli
3	1553.5	91.5	578	16	US-10-672-399-2	Sequence 2, Appli
4	1508.5	88.9	583	10	US-09-902-939-4	Sequence 4, Appli
5	1363	80.3	376	16	US-10-672-399-4	Sequence 4, Appli
6	1298	76.5	360	16	US-10-672-399-6	Sequence 6, Appli
7	965.5	56.9	587	13	US-10-042-523-3	Sequence 3, Appli
8	965.5	56.9	588	15	US-10-309-560-10	Sequence 10, Appli
9	935.5	55.1	554	10	US-09-902-939-6	Sequence 6, Appli
10	935	55.1	552	10	US-09-902-939-5	Sequence 5, Appli
11	935	55.1	552	14	US-10-262-526-4	Sequence 4, Appli
12	934	55.0	552	14	US-10-262-526-2	Sequence 2, Appli
13	928	54.7	553	15	US-10-295-027-370	Sequence 370, App


```

; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 583
; TYPE: PRT
; ORGANISM: mouse
US-03-902-939-4

Query Match      88.9%; Score 1508.5; DB 10; Length 583;
Best Local Similarity 89.9%; Pred. No. 1.8e-148;
Matches 293; Conservative 2; Mismatches 6; Indels 25; Gaps 3;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAAR-----GPLDAATARSVALTISAYQEDPAYLRQCL 55
Db 60 AFLSAHLVAQSLFAYLEHRRVAAAARSLAKGPLDAATARSVALTISAYQEDPAYLRQCL 119
Qy 56 ASARALLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNVHQPWEPA-A 114
Db 120 TSARALLYPHTRLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNVHQPWEPA 179
Qy 115 AGAVGAGAYREVEAEDPGLRAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGSDVDYVQ 174
Db 180 TGAVGEGAYREVEAEDPGLRAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGSDVDYVQ 239
Qy 175 VCDSDTRLDPMALLELRVLDDEPRVGVAGGVDVRLINPLDSWVSLSLRYWVAFNVERA 234
Db 240 VCDSDTRLDPMALLELRVLDDEPRVGVAGGVDVRLINPLDSWVSLSLRYWVAFNVERA 299
Qy 235 CQSFVHCVSCISGPIGLYRNLLQOFLFLEAWYNQKFLGTHCTFGDDRHLTNRLMSNGYATK 294
Db 300 CQSFVHCVSCISGPIGLYRNLLQOFLFLEAWYNQKFLGTHCTFGDDRHLTNRLMSNGYATK 359
Qy 295 -----ABGTRWS 301
Db 360 YTSRRCYSETPSSFLRWLSQOTRWS 385

RESULT 5
US-10-672-399-4
; Sequence 4, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-672-399-4

Query Match      80.3%; Score 1363; DB 16; Length 376;
Best Local Similarity 83.4%; Pred. No. 1.6e-133;
Matches 267; Conservative 0; Mismatches 1; Indels 52; Gaps 1;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Qy 61 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNVHQPWEPAAGAVGA 120
Db 121 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNVHQPWEPAAGAVGA 180
Qy 121 GAYREVEAEDPGLRAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGSDVDYVQCDSDT 180
Db 181 GAYREVEAEDPGLRAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGSDVDYVQCDSDT 240
Qy 181 GAYREVEAEDPGLRAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGSDVDYVQCDSDT 240
Db 240 GAYREVEAEDPGLRAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGSDVDYVQCDSDT 300

Query Match      76.5%; Score 1298; DB 16; Length 360;
Best Local Similarity 95.1%; Pred. No. 9.7e-127;
Matches 252; Conservative 3; Mismatches 6; Indels 4; Gaps 1;

Qy 1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 60
Db 61 AFLSAHLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARA 120
Qy 61 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNVHQPWEPAAGAVGA 120
Db 121 LLYPRARLRVLMVVDGNRAEDLYVDMFREVFADEDPATYVWDGNVHQPWEPAAGAVGA 180
Qy 121 GAYREVEAEDPGLRAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGSDVDYVQCDSDT 180
Db 181 GAYREVEAEDPGLRAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGSDVDYVQCDSDT 240
Qy 181 GAYREVEAEDPGLRAVEALVTRRCVCVQAQRWGKREVMYTAFAKALGSDVDYVQCDSDT 240
Db 241 RLDPMALLELRVLDDEPRVGVAGGVDVRLINPLDSWVSLSLRYWVAFNVERA 300

RESULT 7
US-10-042-523-3
; Sequence 3, Application US/10042523
; Publication No. US20020151026A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; CORRESPONDENCE ADDRESS:
; NUMBER OF SEQUENCES: 4
; ADDRESS: Two Millitia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
```

```
;
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/042,523
; FILING DATE: 19-Oct-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 587 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-042-523-3

Query Match          56.9%; Score 965.5; DB 13; Length 587;
Best Local Similarity 60.5%; Pred. No. 1.1e-91;
Matches 184; Conservative 47; Mismatches 62; Indels 11; Gaps 4;

Qy      6 HLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPR 65
Db      72 HLMQSLFAPLEIRRV---NKSELPKSPKTKTVALTIAGYQENPEYLIKLESCKYVYKPK 128

Qy      66 ARLRVLMVVDGNRAEDLYMVMDFREVFADDPATYVWDGNTHQWPWEPAAGAGAGAYRE 125
Db      129 DKLKILVIDGNTDDAYMMEMFKDVFHGEDVGTVMKGNHYHTVKKPE---ETNKGSCPE 185

Qy      126 VEA---EDPGRLAVEALVTRRCVCVAORMGKREVMYTAFAKLGDSVDYVQVCDSDTRL 182
Db      186 VSKPLNEDEGINMVEELVRNKRVCIMQMGKREVMYTAFAQIGTSVDYVQVCDSDTKL 245

Qy      183 DPMALLELRVLDEDPRVGAVGGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFHCY 242
Db      246 DELATVEMVKLESNDMYGAVGGDVRILNPPYDSFISFMSSRLRYWAFNVERACQSYFDCV 305

Qy      243 SCISGPLGLYRNLLQOFLBANYNQKFLGTHCTFGDDRHLTNRLMSMGYATK--ABGTRW 300
Db      306 SCISGPLGMRYNNILQVFLBANYRQKFLGTCTGLGDDRHLTNRLVLSMGYRTKYTHKSRAP 365

Qy      301 SGTP 304
Db      366 SETP 369

RESULT 8
US-10-309-560-10
; Sequence 10, Application US/10309560
; Publication No. US20030235893A1
; GENERAL INFORMATION:
; APPLICANT: WEIGEL, PAUL H
; APPLICANT: KUMARI, KSHAMA
; TITLE OF INVENTION: HYALURONAN SYNTHASES AND METHODS OF MAKING AND USING SAME
; FILE REFERENCE: 35541.082
; CURRENT APPLICATION NUMBER: US/10/309,560
; PRIOR FILING DATE: 2002-12-03
; PRIOR APPLICATION NUMBER: 60/336,105
; PRIOR FILING DATE: 2001-12-03
; NUMBER OF SEQ ID NOS: 92
```

```
;
;
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 588
; TYPE: PRT
; ORGANISM: Xenopus laevis
US-10-309-560-10

Query Match          56.9%; Score 965.5; DB 15; Length 588;
Best Local Similarity 60.5%; Pred. No. 1.1e-91;
Matches 184; Conservative 47; Mismatches 62; Indels 11; Gaps 4;

Qy      6 HLVAQSLFAYLEHRRVAAAARGPLDAATARSVALTISAYQEDPAYLRQCLASARALLYPR 65
Db      72 HLMQSLFAPLEIRRV---NKSELPKSPKTKTVALTIAGYQENPEYLIKLESCKYVYKPK 128

Qy      66 ARLRVLMVVDGNRAEDLYMVMDFREVFADDPATYVWDGNTHQWPWEPAAGAGAGAYRE 125
Db      129 DKLKILVIDGNTDDAYMMEMFKDVFHGEDVGTVMKGNHYHTVKKPE---ETNKGSCPE 185

Qy      126 VEA---EDPGRLAVEALVTRRCVCVAORMGKREVMYTAFAKLGDSVDYVQVCDSDTRL 182
Db      186 VSKPLNEDEGINMVEELVRNKRVCIMQMGKREVMYTAFAQIGTSVDYVQVCDSDTKL 245

Qy      183 DPMALLELRVLDEDPRVGAVGGDVRILNPLDSWVSFLSSRLRYWAFNVERACQSYFHCY 242
Db      246 DELATVEMVKLESNDMYGAVGGDVRILNPPYDSFISFMSSRLRYWAFNVERACQSYFDCV 305

Qy      243 SCISGPLGLYRNLLQOFLBANYNQKFLGTHCTFGDDRHLTNRLMSMGYATK--ABGTRW 300
Db      306 SCISGPLGMRYNNILQVFLBANYRQKFLGTCTGLGDDRHLTNRLVLSMGYRTKYTHKSRAP 365

Qy      301 SGTP 304
Db      366 SETP 369

RESULT 9
US-09-902-939-6
; Sequence 6, Application US/09902939
; Publication No. US20030087850A1
; GENERAL INFORMATION:
; APPLICANT: Philip DeHazy
; APPLICANT: William Chen
; TITLE OF INVENTION: GENE THERAPY FOR DRY EYE SYNDROME
; FILE REFERENCE: 2055/OH020-USO
; CURRENT APPLICATION NUMBER: US/09/902,939
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 554
; TYPE: PRT
; ORGANISM: mouse
US-09-902-939-6

Query Match          55.1%; Score 935.5; DB 10; Length 554;
Best Local Similarity 57.1%; Pred. No. 1.4e-88;
Matches 185; Conservative 41; Mismatches 61; Indels 37; Gaps 6;

Qy      1 AFLSAHLVAQSLFAYLEHRRVAAAARGPLD---AATARSVALTISAYQEDPAYLRQCLAS 57
Db      50 AILGLHLLIQSLFAPLEHRRVRAGR-PLKLHCSQRSRSVALCIAAQEDPEYLRKCLRS 108

Qy      58 ARALLYPRARLRVLMVVDGNRAEDLYMVMDFREVF-ADDPATYVWDGNTHQWPWEPAAG 116
Db      109 AQRIAPF--NLKVMVMVDGNRQEDTYMLDIFHEVLGTEQAGFVFMRSNFHEAGEGTEA 166

Qy      117 AVGAGAYREVEADPGRLAVEALVTRRCVCVAORMGKREVMYTAFAKLGDSVDYVQVC 176
Db      167 SLOEGMER-----VRVVMVASTFSCIMQMGKREVMYTAFAKLGNSVDYIQVC 215

Qy      177 DSDTRLDPMALLELRVLDEDPRVGAVGGDVRILNPLDSWVSFLSSRLRYWAFNVERACQ 236
Db      177 DSDTRLDPMALLELRVLDEDPRVGAVGGDVRILNPLDSWVSFLSSRLRYWAFNVERACQ 236
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 552
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-262-526-4

Query Match      55.1%; Score 935; DB 14; Length 552;
Best Local Similarity 55.0%; Pred.No.1.6e-88;
Matches 176; Conservative 50; Mismatches 60; Indels 34; Gaps 5;

QY    1 AFLSAHLVAQSFLAYLEHRRVAAARGLDAAATARSVALTISAYQEDPAYLRQC LASARA 60
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db     51 AFLASHLIIOSLFALFLEHRKWKKSLETPI--KUNKTVALCIAAYQEDPYLRKCLQSVKR 108
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::

QY    61 LLYPRRLRLVMVDGNNRAEDLVMDMFRVFADDEDPATYYWBDGNVHQHPPEAAAAGAVGA 120
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db     109 LTYP--GIKVMMWDGNSDDDLWMYDFSEVIGRDKSATYIKNPNFHEK-----GP 157

QY    121 GAYREVEARDPGRLAVALRTRRCVCVAQRWGKGKREVMTYAPKALGDSVDYVQVCDSDT 180
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db     158 GETEESHKESSQH--VTQLVLNKKSI CIMQKWGGKREVMTYAFRALGRSYDYVVQVCDSDT 215

QY    181 RLDPMALLELRVLDEDPRGVAGGDVRIINLPDLSMWVSFLSSLRYWVAENVVERACOSYFH 240
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
Db     216 MLDPASSVEVMVKLEEEDPMVGGGVGDVQIILNKYDWSIFLSSRYRWMAFNIERACOSYFG 275

QY    241 CVSCISGPLGIYRNLLIQOFLWAYNQKFLGTHTCTFCGDDRHLTNRMLSMGYATK----- 294
```

```

00      295 -----AEGTRWS 301
01          : ||||
02      336 CLTETPIEYLRWLNQTRWS 355
03
04 RESULT 12
05 US-10-262-526-2
06 ; Sequence 2, Application US/10262526
07 ; Publication No. US20030108531A1
08 ; GENERAL INFORMATION:
09 ; APPLICANT: Hideshige Moriya
10 ; APPLICANT: Yuichi Wada
11 ; TITLE OF INVENTION: GENES FOR TRANSFECTION INTO BONY TISSUES
12 ; FILE REFERENCE: OP1418
13 ; CURRENT APPLICATION NUMBER: US/10/262,526
14 ; CURRENT FILING DATE: 2002-09-30
15 ; PRIOR APPLICATION NUMBER: JP 2001-367091
16 ; PRIOR FILING DATE: 2001-11-30
17 ; NUMBER OF SEQ ID NOS: 4
18 ; SOFTWARE: PatentIn Ver. 2.1
19 ; SEQ ID NO 2
20 ; LENGTH: 552
21 ; TYPE: PRT
22 ; ORGANISM: Homo sapiens
23 US-10-262-526-2

```

[illegible]

QY 295 -----AEGTRWS 301
Db 337 SKCLTETPTKYLRWLNQOTRWS 358

RESULT 15
US-10-011-768B-10
; Sequence 10, Application US/10011768B
; Publication No. US20030073221A1
; GENERAL INFORMATION:
; APPLICANT: Weigel, Paul H.
; APPLICANT: DeAngelis, Paul
; APPLICANT: Kumari, Kshama
; TITLE OF INVENTION: Hyaluronan Synthase Gene and Uses Thereof
; FILE REFERENCE: 3554.011
; CURRENT APPLICATION NUMBER: US/10/011,768B
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 09/178,851
; PRIOR FILING DATE: 1998-10-26
; PRIOR APPLICATION NUMBER: US 60/064,435
; PRIOR FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Paramecium bursaria chorella virus
US-10-011-768B-10

Query Match 24.1%; Score 409; DB 14; Length 568;
Best Local Similarity 32.9%; Pred. No. 1.7e-33;
Matches 103; Conservative 63; Mismatches 105; Indels 42; Gaps 10;

QY 2 FLSAHLVAQSLFAYLEHRR-----VAAAARGPLDAATARSVALTISAYQEDPAYLRQCLAS 57
Db 56 FVFGFLAQVLFSELNRKLRKRWISLRPKGNDV----RLAVIIAGYREDPTFMFOKCLAS 111

QY 58 ARALLYPRARLRVLMVDGNRAEDLYMDMFREVFEDPADPYVWDGNYHQWEPAAAGA 117
Db 112 VRDSYGNV-ARLICVIDGDEDDMRMAVYKAYND-----NIKKP----- 152

QY 118 VGAGAYREVEADPGRLAVALVTRRCVCVAQRWGGKREVMYTAFK--ALGDSVDYVQV 175
Db 153 ----BFVLCESDKEGERIDS--DFSODICVLQPHRGKRECLYTGFLAKMDPSVNAVVL 206

QY 176 CDSDFRLDPMLLELVRLVLDDEPRVGAVGVDYRILNPLDSWVSFLSSLRYVAFNVERAC 235
Db 207 IDSDTVLEKDAILEVYVYPLACDPEIQAVAGECKIWN-TDTLLSLVAVWRYSAFCVERSA 265

QY 236 QSYFHCVCISGFLGYRNLLQGFLEAWYNOKFLGTHCTFGDDRHLTNRLSMG---Y 291
Db 266 QSFRTVQCVGFLGAYKIDIIKEIKDPWISORFLOQKCTYGGDRRLTNEILMRGKKVVF 325

QY 292 ATKAGTRWSGTP 304
Db 326 TPFVAG--WSDSP 336

Search completed: March 11, 2005, 14:44:39
Job time : 72.8182 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 02:13:47 ; Search time 94.4613 Seconds
(without alignments)
10259.279 Million cell updates/sec

Title: US-10-672-399-9

Perfect score: 20

Sequence: 1 cgggcttgcagactactt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_ats.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	1581	6	CQ731079 Sequence
2	20	100.0	2088	9	HSU59269
3	20	100.0	2108	9	HUMHAS
4	20	100.0	2109	9	AY463695
5	20	100.0	2116	6	AR220003 Sequence
6	20	100.0	2117	6	AR137038 Sequence
7	20	100.0	2119	9	BC035837
8	20	100.0	185623	2	AC137058
9	20	100.0	193986	2	AC130783
10	20	100.0	229155	9	AC018755
11	18.4	92.0	30	6	E28458
12	18.4	92.0	517	4	AB017803
13	18.4	92.0	605	11	BV161384
14	18.4	92.0	621	11	BV099322
15	18.4	92.0	621	11	BV160507
16	18.4	92.0	2095	10	AB097568
17	18.4	92.0	2102	6	E13681
18	18.4	92.0	2102	6	E28454
19	18.4	92.0	2102	6	E30971

20	18.4	92.0	2102	6	E34326
21	18.4	92.0	2102	10	MUSHAS
c 22	18.4	92.0	72955	2	AC108651.3
23	18.4	92.0	257325	2	AC116203
c 24	18.4	92.0	293184	2	AC079487
c 25	18	90.0	285648	2	AC129766
c 26	17.4	87.0	1009	5	CR406688
27	17.4	87.0	2006	5	DRU31079
28	17.4	87.0	2060	5	BC071301
29	17.4	87.0	91871	9	AY340073
c 30	17.4	87.0	149872	2	AC151565
31	17.4	87.0	170962	2	AL389927
32	17.4	87.0	175594	9	AL162503
c 33	16.8	84.0	562	11	G75028
c 34	16.8	84.0	642	11	BV051323
35	16.8	84.0	938	11	G75183
c 36	16.8	84.0	30382	3	CEP42811
c 37	16.8	84.0	78835	9	AL390754
c 38	16.8	84.0	84701	2	AC016122
c 39	16.8	84.0	111486	2	AC136894
40	16.8	84.0	140491	8	AP003276
41	16.8	84.0	142680	8	AP004611
c 42	16.8	84.0	199050	2	AC102844
c 43	16.8	84.0	215046	2	AF354168
c 44	16.8	84.0	218800	2	AC094410
c 45	16.8	84.0	226362	2	AC094851

ALIGNMENTS

RESULT 1	CQ731079	Sequence 17013 from Patent WO02068579.	1581 bp	DNA	linear	PAT 03-FEB-2004
LOCUS	CQ731079					
DEFINITION	CQ731079					
ACCESSION	CQ731079.1	GI:42306695				
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						
	Homo sapiens (human)					
	Homo sapiens					
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE						
AUTHORS	Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.					
TITLE	Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof					
JOURNAL	Patent: WO 02068579-A 17013 06-SEP-2002;					
FEATURES	PE Corporation (NY) (US)					
source	Location/Qualifiers					
	1..1581					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
ORIGIN						
Query Match	100.0%;	Score 20;	DB 6;	Length 1581;		
Best Local Similarity	100.0%;	Pred. No. 2.6;				
Matches	20;	Conservative	0;	Mismatches	0;	Gaps 0;
Qy	1	CGGGCTTCTCAGAGCTACTT	20			
Db	727	CGGGCTTCTCAGAGCTACTT	746			
RESULT 2	HSU59269	Human hyaluronan synthase mRNA, complete cds.	2088 bp	mRNA	linear	PRI 24-SEP-1996
LOCUS	HSU59269					
DEFINITION	HSU59269					
ACCESSION	U59269					
VERSION	U59269.1	GI:1556464				
KEYWORDS						
SOURCE						
	Homo sapiens (human)					

ORGANISM	Homo sapiens	source	1. .2108
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		/organism="Homo sapiens"
AUTHORS	1 (bases 1 to 2088)		/mol_type="mRNA"
TITLE	Shyjan,A.M., Heldin,P., Butcher,E.C., Yoshino,T. and Briskin,M.J.		/db_xref="taxon:9606"
JOURNAL	Functional cloning of the cDNA for a human hyaluronan synthase		/tissue_type="brain"
MEDLINE	J. Biol. Chem. 271 (38), 23395-23399 (1996)		/clone_lib="lgt11"
PUBMED	96394438	gene	/dev_stage="fetal"
REFERENCE	2 (bases 1 to 2088)		1. .2108
AUTHORS	Briskin,M.J. and Shyjan,A.M.	CDS	/gene="hHAS"
TITLE	Direct Submission		149. .1780
JOURNAL	Submitted (24-MAY-1996) LeukoSite Inc., 215 First Street, Cambridge, MA 02142, USA		/gene="hHAS"
FEATURES	Location/Qualifiers		/codon_start=1
source	1. .2088		/product="hyaluronan synthase"
	/organism="Homo sapiens"		/protein_id="AAC50706.1"
	/mol_type="mRNA"		/db_xref="taxon:9606"
CDS	36. .1772		/db_xref="taxon:9606"
	/codon_start=1		/product="hyaluronan synthase"
	/product="hyaluronan synthase"		/protein_id="AAC50706.1"
	/db_xref="taxon:9606"		/db_xref="taxon:9606"
	/translation="MRQODAPKPTPAARRCSGLARRVLTIAFALLILGLMTWYAAGV PLASDRYGLLAGFLYGAFLSAHLVAQSLFAYLEHRRVAAARGLDARTARSVALTIS AYQEDPAYLRQCLASARALLTPRAKRLVIMVVDGNRAEDLYVDMFREVFADEDPATY VMDGNYPQWPEAAGAGVAGAYREVEAEPGLRAVEALVETRCVCAQRMGGKREV MYTAFKLDGSDVYVQVCDSDTRLDPMALLSLRVLDSDPRVAGVGGDVRILNPLDSW VSFLSLRYVAFNVRACQSYFHCVCISGFLGLYRNMLLQQLFLEAWYNOKFLGTGHC TFGDDRHLTNMLSGVATKYTSRSCYSETPSSFLRWLSOOTRWSKSYPREWLYNAL WHRHAWMTVEAVVSGLFFPEFVAATVRLFYAGRPWALLVWLLVCVQVALAKAPAA WLRGLRWLLSLVAPLTMCGLLPAKFLVLMQSGMTSGRRKLAAVYVPLPLPLAL WALLLGLGVSVAHEARADSGPSRAAEAYHLAAGAGVYGVWVAMLTLYVWVGVRLL CRRRTGGYRVQV"		/translation="MTWAYAGVPLASDRYGLLAGFLYGAFLSAHLVAQSLFAYLEHRRVAAARGLDARTARSVALTIS RAEDLRYVMDFREVFADEDPATYVMDGNYPQWPEAAGAGVAGAYREVEAEPGLRA VEALVETRCVCAQRMGGKREVMTAFKALGDSVDYVQVCDSDTRLDPMALLSLRV LDEDPRVAGVGGDVRILNPLDSWVSFLSLRYVAFNVRACQSYFHCVCISGFLGL YRNMLLQQLFLEAWYNOKFLGTGHCFTGDDRHLTNMLSGVATKYTSRSCYSETPSS FLRWLSOOTRWSKSYFREWLYNALWHRHAWMTVEAVVSGLFFPEFVAATVRLFYAGR PALLWLLVCVQVALAKAPAAWLRGLRWLLSLVAPLTMCGLLPAKFLVLMQSGMTSGRRKLAAVYVPLPLPLAL SGMTSGRRKLAAVYVPLPLALWALLLGLGVSVAHEARADSGPSRAAEAYHLAAGAGVYGVWVAMLTLYVWVGVRLL CRRRTGGYRVQV"
ORIGIN		Query Match	100.0%; Score 20; DB 9; Length 2108;
		Best Local Similarity	100.0%; Pred. No. 2.6;
		Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CGGGCTTGTACAGCTACTT 20		
Db	920 CGGGCTTGTACAGCTACTT 939		
RESULT 4			
AY463695			
LOCUS	AY463695	2109 bp	mRNA linear PRI 07-DEC-2003
DEFINITION	Papio anubis hyaluronan synthase (HAS1) mRNA, complete cds.		
ACCESSION	AY463695		
VERSION	AY463695.1	GI:38607341	
KEYWORDS			
SOURCE	Papio anubis (olive baboon)		
ORGANISM	Papio anubis		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Cercopitheidae; Papio.		
REFERENCE	1 (bases 1 to 2109)		
AUTHORS	Martinez-Duncker,I., Oriol,R. and Mollicone,R.		
TITLE	Evolution of the hyaluronan, nodulation C, chitin and cellulose synthases: a superfamily of cell-wall associated carbohydrate polymerizing enzymes		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2109)		
AUTHORS	Martinez-Duncker,I., Oriol,R. and Mollicone,R.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-NOV-2003) Unite 504 Glycobiology Et Signalisation Cellulaire, Inserm, 16 Avenue Paul Vaillant Couturier, Villejuif 94807, France		
FEATURES	Location/Qualifiers		
source	1. .2109		
	/organism="Papio anubis"		
	/mol_type="mRNA"		
	/db_xref="taxon:9555"		
gene	1. .2109		
	/gene="HAS1"		
CDS	36. .1787		
	/gene="HAS1"		
	/note="cell wall associated protein; produces hyaluronic acid chains"		
	/codon_start=1		
	/product="hyaluronan synthase"		

/protein_id="AAR25554.1"
/db_xref="GI:38607342"
/translation="MTQDTPKPTPAARRCSGLARRVLTITAFALLILGLMTWAYAGV
PLATSDRYGLLAFGLYGAFLSAHLLAQSLPAYLEHRRVAAARAAARGRLDAATARSV
ALTSAYQEDPAYLRQCLVARSALLYPPARLVLMVDGNRPEDLIYVDMREVPFAD
DPATVWDGNTQHPREPAAVGVGATREVEADPGRLAVALVTRRCVCVCAQRWG
GKREVVYTFKALGDSVDYQVCDSDTRLDPMALLELVQLDLEPRVAGVGDVRLIN
PLDSVSEFLSLRYVWAFNVERACQSYFHCVCISGPIGLYNNILQQFLAWYVYKQF
LGTHCTFGDRLHTRMLSMGYATKYTSRCYCSFSLRWLSQQTRWSKSYFREW
LYNALWHRHHAMTYEAVSGLPFPFVAATVLRFLFYAGRPWALLWLLVCQGVALAK
AAFAALRGCLRMVLLSYAPLYMCGLLPAKFLALVYTNQSGWGTSGRRKLAANYPL
LPLALWALLLGLLVRSVAHEARADWSGFSRAEYHLAAGAGATVGYVWVWLTLYWV
GVRRLCRRTTGGYRVQV"

ORIGIN
Query Match 100.0%; Score 20; DB 9; Length 2109;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTGACAGCTACTT 20
|||||
Db 927 CGGGCTTGTGACAGCTACTT 946

RESULT 5
LOCUS AR220003 2116 bp mRNA linear PAT 26-SEP-2002
DEFINITION Sequence 1 from patent US 6423514.
ACCESSION AR220003
VERSION AR220003.1 GI:23324403
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 2116)
AUTHORS Briskin,M.J.
TITLE Mammalian hyaluronan synthases, nucleic acids and uses thereof
JOURNAL Patent: US 6423514-A 1 23-JUL-2002;
FEATURES
1..2116
source
/organism="unknown"
/mol_type="mRNA"

ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 2116;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTGACAGCTACTT 20
|||||
Db 912 CGGGCTTGTGACAGCTACTT 931

RESULT 6
LOCUS AR137038 2117 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6162908.
ACCESSION AR137038
VERSION AR137038.1 GI:14478288
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE
1 (bases 1 to 2117)
AUTHORS Itano,N. and Kimata,K.
TITLE Polypeptide of human-origin hyaluronate synthetase and DNA encoding the same
JOURNAL Patent: US 6162908-A 1 19-DEC-2000;
FEATURES
1..2117
source
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 2117;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTGACAGCTACTT 20
|||||
Db 920 CGGGCTTGTGACAGCTACTT 939

RESULT 7
LOCUS BC035837 2119 bp mRNA linear PRI 30-JUN-2004
DEFINITION Homo sapiens hyaluronan synthase 1, mRNA (cDNA clone MGC:46218
IMAGE:5589083), complete cds.
ACCESSION BC035837
VERSION BC035837.1 GI:23243101
KEYWORDS
SOURCE MGC.
ORGANISM Homo sapiens (human)

REFERENCE
1 (bases 1 to 2119)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,P.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S.,
Carninci,P., Frange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Maman,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Maman,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywicki,M.I., Skalek,U., Smalace,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and de Marra,M.A.,
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 2119)
Direct Submission
Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov
Tissue Procurement: Invitrogen
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgri.nih.gov

Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgonov,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 79 Row: f Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4504338.

FEATURES
source

Location/Qualifiers
1..2119
/organism="Homo sapiens"
/no_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:46218 IMAGE:5589083"
/tissue type="Ovary, pooled from 3 adults"
/clone_lib="NIH MGC_125"
/lab host="DH10B"
/note="Vector: pCMV-SPORT6"

gene

1..2119
/gene="HAS1"
/note="synonym: HAS"
/db_xref="LocustID:3036"
/db_xref="MIM:601463"
42..1775
/gene="HAS1"

CDS

/codon_start=1
/product="hyaluronan synthase 1"
/protein_id="AAH35837.1"
/db_xref="GI:23243102"
/db_xref="LocustID:3036"
/db_xref="MIM:601463"
/translation="MRQAPKPTPAARCSGLARRVLTAFALLILGLMTWYAAAGVP
LASDRYGLAFLGLYAGLAHLVAQSLPAYLEHRRVAAARGLDAAARSVALTISA
YQEDPAYLRQCLASARALLYPRARLVLMVDGNRAEDLYMDVMPREVFADDPATYV
WDGNYHQPEAPAGACAGAYREAEADPGRLAVALRTRCVCVAORWGGKREVM
YTFKALGDSDHYQVDCSDTRLDPMALLELVRLVLEDDPRVAGVGGDVRIINPLDSWV
SFLSLRLVAPNVERACQSFHCVSCISGELGLYRNLLQQLFLEAWYQKFLGTHTC
FGDHLNRLMSGYATKYTSRSRCYSETSPSLRWLSQOTRWSKSYFRELWYNALW
WRHAWNTYEAIVGSPFPFVAATVRLFLFYAGRPWALLVLLCVQVALAKAFAAW
LGLGLRVLLSLYAPLYWGLLPKFLALVTMNSQWGTSGRRKLAAHYVPLLPALW
ALLLGLGLRVVLAHEARDWSPSPRAAEAYHLAAGAYGVYVWMLTYWVGVRRLC
RRRTGGYRVQV"

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 2119;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTGACGACTACTT 20

Db 915 CGGGCTTGTGACGACTACTT 934

RESULT 8

AC137058/c
LOCUS
DEFINITION AC137058 185623 bp DNA linear HTG 21-FEB-2003
Papio anubis clone RP41-126M5, WORKING DRAFT SEQUENCE, 10 ordered
pieces.
ACCESSION AC137058
VERSION AC137058.3 GI:28460766
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Papio anubis (olive baboon)
ORGANISM Papio anubis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopithecinae; Papio.
1 (bases 1 to 185623)
Akter.N., Antonellis.A., Ayele.K., Beckstrom-Sternberg.S.M.,
Benjamin.B., Blakesley.R.W., Bouffard.G.G., Brinkley.C., Brookes.S.,
Cariaga.K., Coleman.B., Engle.J., Granite.S., Guan.X., Gupta.J.,
Hachichi.P., Han.J., Hansen.N., Ho.S.-L., Idol.J.R., Karlins.E.,
Laric.P., Lee-Lin.S.-Q., Legaspi.R., Maduro.Q.I., Maduro.V.B.,
Margulies.E.H., Mastello.C., Maskeri.B., McDowell.J.,
Paguirigan.C., Pearson.R., Portnoy.M.E., Prasad.A.,
Reddix-Dugue.N., Schandler.K., Schueler.M.G., Sison.C.,
Stantripop.S., Thomas.J.W., Thomas.P.J., Touchman.J.W., Vogt.J.L.,
Wetherby.K.D., Wiggins.L., Young.A. and Green.E.D.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 185623)
Green,E.D.
Direct Submission
Submitted (15-NOV-2002) NIH Intramural Sequencing Center, 8717
Groveont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 185623)
Green,E.D.
Direct Submission
Submitted (21-FEB-2003) NIH Intramural Sequencing Center, 8717
Groveont Circle, Gaithersburg, MD 20877, USA
On Feb 21, 2003 this sequence version replaced gi:27476124.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: <http://www.nisc.nih.gov>
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: dtj
Center clone name: 126M05

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics -----

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 183052 bases at least Q40
Consensus quality: 183851 bases at least Q30
Consensus quality: 184434 bases at least Q20
Insert size: 160000; agarose-fp
Insert size: 184723; sum-of-contigs
Quality coverage: 14.13x in Q20 bases; agarose-fp
Quality coverage: 12.24x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 9984: contig of 9984 bp in length

* 9985 10084: gap of unknown length

* 10085 49805: contig of 39721 bp in length

* 49806 49905: gap of unknown length

* 49906 66936: contig of 17031 bp in length

* 66937 67036: gap of unknown length

* 67037 83604: contig of 16568 bp in length

* 83605 83704: gap of unknown length

* 83705 108246: contig of 24542 bp in length

* 108247 108346: gap of unknown length

* 108347 112021: contig of 3675 bp in length

* 112022 112121: gap of unknown length

* 112122 134817: contig of 22696 bp in length

* 134818 134917: gap of unknown length

* 134918 147290: contig of 12373 bp in length

* 147291 147390: gap of unknown length

* 147391 185337: contig of 37947 bp in length

* 185338 185437: gap of unknown length

FEATURES * 185438 185623: contig of 186 bp in length.

source
 1. 185623
 /organism="Papio anubis"
 /mol_type="genomic DNA"
 /db_xref="taxon:9555"
 /clone="RP41-126M5"
 /clone_lib="RP41"
 1. 15263
 /note="clone overlaps with GenBank Accession Number AC130273 clone RP41-30706 (center project name dsu)"
 misc_feature
 1. 9984
 /note="assembly_fragment
 clone_end:T7
 vector_side:left"
 misc_feature
 10085. .49805
 /note="assembly_fragment"
 misc_feature
 49906. 66936
 /note="assembly_fragment"
 misc_feature
 67037. .83604
 /note="assembly_fragment"
 misc_feature
 83705. 108246
 /note="assembly_fragment"
 misc_feature
 108347. 112021
 /note="assembly_fragment"
 misc_feature
 112122. 134817
 /note="assembly_fragment"
 misc_feature
 134918. 147290
 /note="assembly_fragment"
 misc_feature
 147391. 185337
 /note="assembly_fragment"
 misc_feature
 185438. 185623
 /note="assembly_fragment
 clone_end:SP6
 vector_side:right"

ORIGIN

Query Match 100.0%; Score 20; DB 2; Length 185623;
 Best Local Similarity 100.0%; Pred. No. 3.4;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCTTCTCAGACTACTT 20
 |||||
 Db 135015 CGGCTTCTCAGACTACTT 134996

RESULT 9
 AC130783/c
 LOCUS AC130783 193986 bp DNA linear HTG 19-NOV-2002
 DEFINITION Pan troglodytes clone CH251-426A12, WORKING DRAFT SEQUENCE, 8
 ordered pieces.
 ACCESSION AC130783
 VERSION AC130783.2 GI:25100968
 KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 1 (bases 1 to 193986)
 Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
 Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
 Cariaga,K., Coleman,B., Engle,J., Granite,S., Guan,X., Gupta,J.,
 Haghighi,P., Han,J., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E.,
 Latic,P., Lee-Lin,S.-Q., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
 Margulies,E.H., Mastello,C., Maskeri,B., McDowell,J.,
 Paquirigan,C., Pearson,R., Portnoy,M.E., Prasad,A.,
 Reddix-Dugue,N., Schandler,K., Schueler,M.G., Sison,C.,
 Stantripop,S., Thomas,J.W., Thomas,P.J., Touchman,J.W., Vogt,J.L.,
 Wetherby,K.D., Wiggins,L., Young,A. and Green,E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 193986)
 Green,E.D.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Direct Submission
 Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 193986)
 Green,E.D.
 Direct Submission
 Submitted (19-NOV-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 On Nov 19, 2002 this sequence version replaced gi:22218453.
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc.zoo@nhgri.nih.gov
 ----- Project Information
 Center project name: dxid
 Center clone name: 426A12

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 151764 bases at least Q40
 Consensus quality: 192510 bases at least Q30
 Consensus quality: 192998 bases at least Q20
 Insert size: 184000; agarose-fp
 Quality coverage: 11.49x in Q20 bases; agarose-fp
 Quality coverage: 10.94x in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently consists of 8 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced
 * by the finished sequence as soon as it is available and the accession number will be preserved.

* 11512: contig of 11512 bp in length
 * 11513 11612: gap of unknown length
 * 11613 59488: contig of 47876 bp in length
 * 59489 59588: gap of unknown length
 * 59589 103266: contig of 43678 bp in length
 * 103267 103366: gap of unknown length
 * 103367 127885: contig of 24519 bp in length
 * 127886 127985: gap of unknown length
 * 127986 129692: contig of 1707 bp in length
 * 129693 129792: gap of unknown length
 * 131077 131177: contig of 1285 bp in length
 * 131078 131178: gap of unknown length
 * 131178 174663: contig of 43486 bp in length
 * 174664 174763: gap of unknown length
 * 174764 193986: contig of 19223 bp in length.

FEATURES
 source
 1. 193986
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="CH251-426A12"
 /clone_lib="CH251"


```
repeat_region /rpt_family="LMB"
6936..7295
repeat_region /rpt_family="MER33"
7700..7785
repeat_region /rpt_family="L2"
complement(9216..9434)
repeat_region /rpt_family="LMB"
complement(9435..9745)
repeat_region /rpt_family="AluY"
complement(9746..9815)
repeat_region /rpt_family="LMB"
complement(10009..10303)
repeat_region /rpt_family="AluSp"
11783..12637
repeat_region /rpt_family="LMB2"
12180..12330
STS /gene="ZNF175"
/standard_name="SHGC-32377"
13436..13619
repeat_region /rpt_family="MIR"
16068..16112
repeat_region /rpt_family="(TA)n"
complement(16113..16329)
repeat_region /rpt_family="MER46A"
16330..16390
repeat_region /rpt_family="(TA)n"
16403..16439
repeat_region /rpt_family="(TATG)n"
16793..17013
repeat_region /rpt_family="MER58A"
17024..17052
repeat_region /rpt_family="(CA)n"
17311..17545
repeat_region /rpt_family="MIR"
complement(18683..18896)
repeat_region /rpt_family="MIR"
complement(18915..19183)
repeat_region /rpt_family="AluY"
19336..19599
repeat_region /rpt_family="MIR"
19811..19928
repeat_region /rpt_family="GA-rich"
22181..23081
repeat_region /rpt_family="L2"
complement(23692..23901)
repeat_region /rpt_family="L2"
complement(23902..24201)
repeat_region /rpt_family="AluSg1"
complement(24202..24651)
repeat_region /rpt_family="L2"
26136..26434
repeat_region /rpt_family="LTR12"
26435..26525
repeat_region /rpt_family="LTR30"
26526..26643
repeat_region /rpt_family="LTR12"
complement(26686..26884)
repeat_region /rpt_family="AluSg/x"
27065..27240
repeat_region /rpt_family="MIR"
complement(27238..27554)
repeat_region /rpt_family="L2"
28406..28433
repeat_region /rpt_family="GC-rich"
28641..28678
repeat_region /rpt_family="GC-rich"
complement(28997..29259)
repeat_region /rpt_family="MSR1"
complement(30219..30489)
repeat_region /rpt_family="L2"
30753..30882
repeat_region /rpt_family="FLAM_C"
31138..31416

/rpt_family="AluY"
complement(31547..31716)
repeat_region /rpt_family="MER3"
32041..32112
repeat_region /rpt_family="LMBc"
32117..32516
repeat_region /rpt_family="MLT2E"
35056..35118
repeat_region /rpt_family="MIR"
complement(35578..35636)
repeat_region /rpt_family="MER5A"
complement(35690..35820)
repeat_region /rpt_family="MIR"
35872..35902
repeat_region /rpt_family="(TG)n"
complement(35906..36209)
repeat_region /rpt_family="AluJo"

Query Match 100.0%; Score 20; DB 9; Length 229155;
Best Local Similarity 100.0%; Pred.No. 3.5;
Matches: 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTACAGACTACTT 20
Db 144168 CGGGCTTGTACAGACTACTT 144149

RESULT 11
E28458
LOCUS E28458 30 bp DNA linear PAT 18-JUN-2001
DEFINITION Hyaluronate synthase promoter DNA.
ACCESSION E28458
VERSION E28458.1 GI:13018350
KEYWORDS JP 1999196875-A/6.
SOURCE unclassified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 30)
AUTHORS Voichi,Y., Naoki,I. and Koji,K.
TITLE Hyaluronate synthase promoter DNA
JOURNAL Patent: JP 1999196875-A 6 27-JUL-1999;
SEIRAGAKU KOGYO CO LTD
COMMENT OS Unidentified
PN JP 1999196875-A/6
PD 27-JUL-1999
PF 14-JAN-1998 JP 1998006191
PR YOICHI YAMADA,NAOKI IPANO,KOJI KIMATA
PC C12N15/09,C12N9/00,C12Q1/68//C12N15/09,C12R1:91),C12N15/00,
CC (C12N15/00,C12R1:91)
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..30
FT Location/Qualifiers
1..30
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

FEATURES
source
1..30
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 92.0%; Score 18.4; DB 6; Length 30;
Best Local Similarity 95.0%; Pred.No. 19;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGGCTTGTACAGACTACTT 20
Db 7 CGAGCTTGTACAGACTACTT 26

RESULT 12
AB017803
```

LOCUS AB017803 517 bp mRNA linear MAM 02-APR-1999
 DEFINITION Bos taurus mRNA for hyaluronan synthase 1, partial cds.
 ACCESSION AB017803
 VERSION AB017803.1 GI:4586933
 KEYWORDS HAS1; hyaluronan synthase 1.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovinae; Bos.
 REFERENCE 1 (bases 1 to 517)
 AUTHORS Yamashita,H., Usui,T. and Suzuki,K.
 TITLE HAS1
 JOURNAL Published Only in Database (1999)
 REFERENCE 2 (bases 1 to 517)
 AUTHORS Yamashita,H., Usui,T. and Suzuki,K.
 TITLE Direct Submission
 JOURNAL Submitted (21-SEP-1998) Hidetoshi Yamashita, Faculty of Medicine,
 University of Tokyo, Department of Ophthalmology; Hongo 7-3-1,
 Bunkyo-ku, Tokyo 113-8655, Japan
 (E-mail:hyama-cky@umin.u-tokyo.ac.jp, Tel: +81-3-3815-5411 (ex.3494),
 Fax: +81-3-3817-0798)

FEATURES

source

1..517
 /location/Qualifiers
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 1..517
 /gene="HAS1"
 <1..>517
 /gene="HAS1"
 /codon_start=1
 /product="hyaluronan synthase 1"
 /protein_id="BAA76546.1"
 /db_xref="GI:4586934"
 /translation="RACQSYFHCVCISGPGLYRNLLQOFLEAWYKQFLGTHCTF
 GDRRLNRMLSMGVATKYTSRRCVSETPSSFLWLSQOTRWKSYPREWLYNALW
 HRHWAMTYEAVSGLFFPFFVAAVTLRLFYAGRPWALLWLLVCQGVALLAKAAPAWL
 RGLRMVLLSLY"

ORIGIN

Query Match 92.0%; Score 18.4; DB 4; Length 517;
 Best Local Similarity 95.0%; Pred. No. 22;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCAGAGTACTT 20

||||| ||||| ||||| ||||| |||||
 Db 1 CGGGCTTGCCAGAGTACTT 20

RESULT 13
 BV161384
 LOCUS BV161384 605 bp DNA linear STS 15-MAY-2004
 DEFINITION RPAMMSQ0038215 Roche Palo Alto Mus musculus STS genomic, sequence
 tagged site.
 ACCESSION BV161384
 VERSION BV161384.1 GI:47265166
 KEYWORDS STS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 605)
 AUTHORS Usuka,J., Liao,G., Cheng,J., Nguyen,A., Bach,C., Puech,A.,
 McPherson,J.D., Foerzler,D. and Peltz,G.
 TITLE Mus musculus SNPs
 JOURNAL Unpublished (2003)
 COMMENT

Contact: Jonathan Usuka
 Roche Palo Alto Genetics and Genomics Department
 Roche Palo Alto
 3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA
 Tel: 6508555807

FEATURES

source

Email: Jonathan.Usuka@roche.com
 Primer A: No primer submitted with this STS
 Primer B: No primer submitted with this STS.
 Location/Qualifiers
 1..605
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /map="17-12488-11884-CAAA01219574.1.1.20118"
 /clone_lib="Roche Palo Alto"
 /note="SNPs developed from assay sequences derived from 15
 different strains of mice (as of October 1, 2003). Those
 strains include A/J, A/HeJ, 129/Sv, AKR/J, B10.D2-H2/cSnJ,
 BALB/cByJ, BALB/cJ, C3H/HeJ, C57BL/6J, CAST/Ei, DBA/2J,
 MRL/MpJ, NZB/BinJ, NZW/LaC, SPRET/Ei."
 <1..>605

STS

ORIGIN

Query Match 92.0%; Score 18.4; DB 11; Length 605;
 Best Local Similarity 95.0%; Pred. No. 22;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCAGAGTACTT 20

||||| ||||| ||||| ||||| |||||
 Db 440 CGAGCTTGTCAGAGTACTT 459

RESULT 14

BV099322/c

LOCUS

DEFINITION RPAMMSQ0014265 Roche Palo Alto Mus musculus STS genomic, sequence
 tagged site.
 ACCESSION BV099322
 VERSION BV099322.1 GI:40812400
 KEYWORDS STS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 621)
 AUTHORS Usuka,J., Liao,G., Cheng,J., Nguyen,A., Bach,C., Puech,A.,
 McPherson,J.D., Foerzler,D. and Peltz,G.
 TITLE Mus musculus SNPs
 JOURNAL Unpublished (2003)
 COMMENT

Contact: Jonathan Usuka
 Roche Palo Alto Genetics and Genomics Department
 Roche Palo Alto
 3431 Hillview Ave, Mailstop S3-1, Palo Alto, CA 94024, USA
 Tel: 6508555807

Email: Jonathan.Usuka@roche.com

Primer A: No primer submitted with this STS

Primer B: No primer submitted with this STS.

Location/Qualifiers

FEATURES

source

1..621
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /map="17-11968-12587-CAAA01219574.1.1.20118"
 /clone_lib="Roche Palo Alto"
 /note="SNPs developed from assay sequences derived from 15
 different strains of mice (as of October 1, 2003). Those
 strains include A/J, A/HeJ, 129/Sv, AKR/J, B10.D2-H2/cSnJ,
 BALB/cByJ, BALB/cJ, C3H/HeJ, C57BL/6J, CAST/Ei, DBA/2J,
 MRL/MpJ, NZB/BinJ, NZW/LaC, SPRET/Ei."
 <1..>621

STS

ORIGIN

Query Match 92.0%; Score 18.4; DB 11; Length 621;
 Best Local Similarity 95.0%; Pred. No. 22;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCAGAGTACTT 20

Search completed: March 13, 2005, 08:01:38
Job time : 97.4613 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 01:47:21 ; Search time 12.9978 Seconds
(without alignments)
9108.809 Million cell updates/sec

Title: US-10-672-399-9

Perfect score: 20

Sequence: 1 cgggcttgcagagctactt 20

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002as.*

7: Geneseqn2002bs.*

8: Geneseqn2003as.*

9: Geneseqn2003bs.*

10: Geneseqn2003cs.*

11: Geneseqn2003ds.*

12: Geneseqn2004as.*

13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	490	ACH39998	ACH39998 Human foe
2	20	100.0	1737	ADL13691	Adl13691 Osteoarth
3	20	100.0	1737	ADL13692	Adl13692 Osteoarth
4	20	100.0	2087	ADL13694	Adl13694 Osteoarth
5	20	100.0	2088	ADL13690	Adl13690 Osteoarth
6	20	100.0	2116	AAT99541	Aat99541 Human hya
7	20	100.0	2117	AAT96713	Aat96713 Human hya
8	20	100.0	231222	ADL13693	Adl13693 Osteoarth
9	18.4	92.0	30	AAT10866	Aat10866 PCR prime
10	18.4	92.0	1752	ABZ76734	Abz76734 Mouse hya
11	18.4	92.0	1752	AAD59442	Aad59442 Mouse hya
12	18.4	92.0	2102	AAT91855	Aat91855 Mouse hya
13	18.4	92.0	2102	AAZ10862	Aaz10862 Hyaluroa
14	18.4	92.0	2102	AAZ88199	Aaz88199 Mouse hya
15	18.4	92.0	2102	AAA39987	Aaa39987 Murine HA
16	16.4	82.0	25	AAT10869	Aat10869 PCR prime
17	16.4	82.0	2000	ADA73353	Ada73353 Rice gene
18	16.4	82.0	3831	ADF55687	Adf55687 Mouse car
19	16.4	82.0	21410	ACN44668	Acn44668 Mouse gen
20	15.8	79.0	316	ADP93649	Adp93649 Cotton ex

21	15.8	79.0	400	2	AAV78369	Aav78369 Staphyloc
22	15.8	79.0	430	6	ABN22318	Abn22318 Human ORF
23	15.8	79.0	477	4	AAI11803	Aai11803 Probe #17
24	15.8	79.0	477	4	ABA53497	Abas53497 Human foe
25	15.8	79.0	477	4	AAI33119	Aai33119 Probe #18
26	15.8	79.0	477	4	ABA43081	Abas43081 Human bre
27	15.8	79.0	477	4	ABA23262	Abas23262 Probe #17
28	15.8	79.0	477	4	AAK27223	Aak27223 Human bon
29	15.8	79.0	477	4	AAK01768	Aak01768 Human bra
30	15.8	79.0	477	4	ABS26800	Abs26800 Human liv
31	15.8	79.0	477	5	AAI01737	Aai01737 Probe #17
32	15.8	79.0	477	6	ABS01762	Abs01762 Human gen
33	15.8	79.0	582	4	ABA60297	Abas60297 Human foe
34	15.8	79.0	582	4	AAI40181	Aai40181 Probe #88
35	15.8	79.0	582	4	AAK34458	Aak34458 Human bon
36	15.8	79.0	582	4	AAK08575	Aak08575 Human bra
37	15.8	79.0	582	4	ABS34235	Abs34235 Human liv
38	15.8	79.0	1950	8	ACF74273	Acf74273 Staphyloc
39	15.8	79.0	1953	8	ACA20042	ACA20042 Prokaryot
40	15.8	79.0	2111	2	AAT80389	Aat80389 Staphyloc
41	15.8	79.0	9623	2	AAV74477	Aav74477 Staphyloc
42	15.8	79.0	14781	4	AAK66710	Aak66710 Human imm
43	15.8	79.0	42379	12	ADQ97660	Adq97660 Mouse can
44	15.8	79.0	50720	13	ABD33468	Abd33468 Murine ca
45	15.8	79.0	77834	11	ACN44076	Acn44076 Mouse gen

ALIGNMENTS

RESULT 1

ACH39998 standard; cDNA; 490 BP.

XX ACH39998;

XX 13-OCT-2003 (first entry)

XX Human foetal brain cDNA #1365.

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;

XX genome mapping; biodiversity; genetic disorder.

XX Homo sapiens.

XX US2003073623-A1.

XX 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

XX (DRMA/) DRMANAC R T.

XX (LABA/) LABAT I.

XX (STAC/) STACHE-CRAIN B.

XX (DICK/) DICKSON M C.

XX (JONE/) JONES L W.

XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

New polynucleotide sequences obtained from various cDNA libraries, useful as hybridization probes, as oligomers for PCR, for chromosome and gene mapping, in the recombinant production of protein, or in generating antisense DNA or RNA.

Claim 1; SEQ ID NO 27210; 44pp; English.

The invention relates to an isolated polynucleotide comprising any one of 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was determined by the technique of SBH (sequencing by hybridisation). Also

CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocId=20030073623

XX SQ Sequence 490 BP; 91 A; 123 C; 159 G; 108 T; 0 U; 9 Other;

Query Match 100.0%; Score 20; DB 9; Length 490;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTACAGACTACTT 20
|||||
DB 335 CGGGCTTGTACAGACTACTT 354

RESULT 2

ADL13691
ID ADL13691 standard; DNA; 1737 BP.

AC ADL13691;

DT 06-MAY-2004 (first entry)

DE Osteoarthritis-associated polymorphic nucleotide #224.

KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.

OS Homo sapiens.

PN WO2003054166-A2.

PD 03-JUL-2003.

PF 19-DEC-2002; 2002WO-US041225.

PR 20-DEC-2001; 2001US-0342603P.

PA (INCY-) INCYTE GENOMICS INC.

PI Jones KA, Schafer A;

PS WPI; 2003-559141/52.

XX Determining susceptibility of an individual to joint space narrowing,
XX osteophyte development and/or joint pain comprises identifying whether
XX the individual has at least one polymorphism in a polynucleotide encoding
XX a protein.

PS Disclosure; SEQ ID NO 223; 297pp; English.

XX The invention relates to a method of determining susceptibility of an
XX individual to joint space narrowing and/or osteophyte development and/or
XX joint pain comprising identifying whether the individual has at least one
XX polymorphism in a polynucleotide encoding at least one of the protein
XX listed in the specification. The methods, composition and agent are
XX useful for modulating the susceptibility of an individual to joint space
XX narrowing and/or osteophyte development and/or joint pain that is
XX associated with a disease, preferably osteoarthritis. The cell line and
XX the non-human animal are useful for screening for an agent for diagnosing
XX an individual having susceptibility to joint space narrowing and/or

CC osteophyte development and/or joint pain. This sequence corresponds to
CC the polynucleotide encoding a protein listed in the specification. (Note:
CC The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences).

SQ Sequence 1737 BP; 232 A; 571 C; 603 G; 331 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 10; Length 1737;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTACAGACTACTT 20
|||||
DB 877 CGGGCTTGTACAGACTACTT 896

RESULT 3

ADL13692
ID ADL13692 standard; DNA; 1737 BP.

AC ADL13692;

DT 06-MAY-2004 (first entry)

DE Osteoarthritis-associated polymorphic nucleotide #224.

KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
KW joint space narrowing; osteophyte development; joint pain;
KW osteoarthritis; SNP; single nucleotide polymorphism.

OS Homo sapiens.

PN WO2003054166-A2.

PD 03-JUL-2003.

PF 19-DEC-2002; 2002WO-US041225.

PR 20-DEC-2001; 2001US-0342603P.

PA (INCY-) INCYTE GENOMICS INC.

PI Jones KA, Schafer A;

PS WPI; 2003-559141/52.

XX Determining susceptibility of an individual to joint space narrowing,
XX osteophyte development and/or joint pain comprises identifying whether
XX the individual has at least one polymorphism in a polynucleotide encoding
XX a protein.

PS Disclosure; SEQ ID NO 224; 297pp; English.

XX The invention relates to a method of determining susceptibility of an
XX individual to joint space narrowing and/or osteophyte development and/or
XX joint pain comprising identifying whether the individual has at least one
XX polymorphism in a polynucleotide encoding at least one of the protein
XX listed in the specification. The methods, composition and agent are
XX useful for modulating the susceptibility of an individual to joint space
XX narrowing and/or osteophyte development and/or joint pain that is
XX associated with a disease, preferably osteoarthritis. The cell line and
XX the non-human animal are useful for screening for an agent for diagnosing
XX an individual having susceptibility to joint space narrowing and/or
XX osteophyte development and/or joint pain. This sequence corresponds to
XX the polynucleotide encoding a protein listed in the specification. (Note:
XX The sequence data for this patent did not form part of the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences).

SQ Sequence 1737 BP; 232 A; 569 C; 604 G; 332 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 10; Length 1737;

```
Best Local Similarity 100.0%; Pred. No. 3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTTCAGAGTACTT 20
Db 877 CGGGCTTGTTCAGAGTACTT 896

RESULT 4
ADL13694
ID ADL13694 standard; DNA; 2087 BP.
XX AC ADL13694;
XX DT 06-MAY-2004 (first entry)
XX DE Osteoarthritis-associated polymorphic nucleotide #226.
XX KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX KW joint space narrowing; osteophyte development; joint pain;
XX KW osteoarthritis; SNP; single nucleotide polymorphism.
XX OS Homo sapiens.
XX PN WO2003054166-A2.
XX PD 03-JUL-2003.
XX PF 19-DEC-2002; 2002WO-US041225.
XX PR 20-DEC-2001; 2001US-0342603P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Jones KA, Schafer A;
XX DR WPI; 2003-559141/52.
XX PT Determining susceptibility of an individual to joint space narrowing,
XX PT osteophyte development and/or joint pain comprises identifying whether
XX PT the individual has at least one polymorphism in a polynucleotide encoding
XX PS a protein.
XX PS Disclosure; SEQ ID NO 226; 297pp; English.
XX CC The invention relates to a method of determining susceptibility of an
XX CC individual to joint space narrowing and/or osteophyte development and/or
XX CC joint pain comprising identifying whether the individual has at least one
XX CC polymorphism in a polynucleotide encoding at least one of the protein
XX CC listed in the specification. The methods, composition and agent are
XX CC useful for modulating the susceptibility of an individual to joint space
XX CC narrowing and/or osteophyte development and/or joint pain that is
XX CC associated with a disease, preferably osteoarthritis. The cell line and
XX CC the non-human animal are useful for screening for an agent for diagnosing
XX CC an individual having susceptibility to joint space narrowing and/or
XX CC osteophyte development and/or joint pain. This sequence corresponds to
XX CC the polynucleotide encoding a protein listed in the specification. (Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences).
XX SQ Sequence 2087 BP; 302 A; 650 C; 712 G; 423 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 10; Length 2087;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTTCAGAGTACTT 20
Db 912 CGGGCTTGTTCAGAGTACTT 931

RESULT 5
ADL13694
ID ADL13694 standard; DNA; 2087 BP.
XX AC ADL13694;
XX DT 06-MAY-2004 (first entry)
XX DE Osteoarthritis-associated polymorphic nucleotide #226.
XX KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX KW joint space narrowing; osteophyte development; joint pain;
XX KW osteoarthritis; SNP; single nucleotide polymorphism.
XX OS Homo sapiens.
XX PN WO2003054166-A2.
XX PD 03-JUL-2003.
XX PF 19-DEC-2002; 2002WO-US041225.
XX PR 20-DEC-2001; 2001US-0342603P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Jones KA, Schafer A;
XX DR WPI; 2003-559141/52.
XX PT Determining susceptibility of an individual to joint space narrowing,
XX PT osteophyte development and/or joint pain comprises identifying whether
XX PT the individual has at least one polymorphism in a polynucleotide encoding
XX PS a protein.
XX PS Disclosure; SEQ ID NO 226; 297pp; English.
XX CC The invention relates to a method of determining susceptibility of an
XX CC individual to joint space narrowing and/or osteophyte development and/or
XX CC joint pain comprising identifying whether the individual has at least one
XX CC polymorphism in a polynucleotide encoding at least one of the protein
XX CC listed in the specification. The methods, composition and agent are
XX CC useful for modulating the susceptibility of an individual to joint space
XX CC narrowing and/or osteophyte development and/or joint pain that is
XX CC associated with a disease, preferably osteoarthritis. The cell line and
XX CC the non-human animal are useful for screening for an agent for diagnosing
XX CC an individual having susceptibility to joint space narrowing and/or
XX CC osteophyte development and/or joint pain. This sequence corresponds to
XX CC the polynucleotide encoding a protein listed in the specification. (Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences).
XX SQ Sequence 2087 BP; 302 A; 650 C; 712 G; 423 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 10; Length 2087;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTTCAGAGTACTT 20
Db 912 CGGGCTTGTTCAGAGTACTT 931

RESULT 6
AAT99541
ID AAT99541 standard; cDNA; 2116 BP.
XX AC AAT99541;
XX DT 21-MAY-1998 (first entry)
XX DE Human hyaluronan synthase cDNA clone 30C.
```

```
ADL13690
ID ADL13690 standard; DNA; 2088 BP.
XX AC ADL13690;
XX DT 06-MAY-2004 (first entry)
XX DE Osteoarthritis-associated polymorphic nucleotide #222.
XX KW ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
XX KW joint space narrowing; osteophyte development; joint pain;
XX KW osteoarthritis; SNP; single nucleotide polymorphism.
XX OS Homo sapiens.
XX PN WO2003054166-A2.
XX PD 03-JUL-2003.
XX PF 19-DEC-2002; 2002WO-US041225.
XX PR 20-DEC-2001; 2001US-0342603P.
XX PA (INCY-) INCYTE GENOMICS INC.
XX PI Jones KA, Schafer A;
XX DR WPI; 2003-559141/52.
XX PT Determining susceptibility of an individual to joint space narrowing,
XX PT osteophyte development and/or joint pain comprises identifying whether
XX PT the individual has at least one polymorphism in a polynucleotide encoding
XX PS a protein.
XX PS Disclosure; SEQ ID NO 222; 297pp; English.
XX CC The invention relates to a method of determining susceptibility of an
XX CC individual to joint space narrowing and/or osteophyte development and/or
XX CC joint pain comprising identifying whether the individual has at least one
XX CC polymorphism in a polynucleotide encoding at least one of the protein
XX CC listed in the specification. The methods, composition and agent are
XX CC useful for modulating the susceptibility of an individual to joint space
XX CC narrowing and/or osteophyte development and/or joint pain that is
XX CC associated with a disease, preferably osteoarthritis. The cell line and
XX CC the non-human animal are useful for screening for an agent for diagnosing
XX CC an individual having susceptibility to joint space narrowing and/or
XX CC osteophyte development and/or joint pain. This sequence corresponds to
XX CC the polynucleotide encoding a protein listed in the specification. (Note:
XX CC The sequence data for this patent did not form part of the printed
XX CC specification but was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences).
XX SQ Sequence 2088 BP; 302 A; 652 C; 711 G; 423 T; 0 U; 0 Other;

Query Match 100.0%; Score 20; DB 10; Length 2088;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTTCAGAGTACTT 20
Db 912 CGGGCTTGTTCAGAGTACTT 931

RESULT 6
AAT99541
ID AAT99541 standard; cDNA; 2116 BP.
XX AC AAT99541;
XX DT 21-MAY-1998 (first entry)
XX DE Human hyaluronan synthase cDNA clone 30C.
```

KW Hyaluronan synthase; HAS gene; human; hyaluronic acid; cell adhesion;
 KW wound healing; vulnery; tissue repair; scar; keloid; therapy; ds.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 36..1772
 FT polyA_signal /*tag= a
 FT 2066..2071 /*tag= b
 XX
 XX WO9740174-A1.
 XX
 XX 30-OCT-1997.
 XX
 XX 17-APR-1997; 97WO-US006350.
 XX
 XX 22-APR-1996; 96US-00635552.
 XX (LEUK-) LEUKOSITE INC.
 PA Brieskin MJ;
 PI WPI; 1997-549359/50.
 DR P-PSDB; AAW26765.
 XX
 XX Human hyaluronan synthase - useful for recombinant production of
 PT hyaluronic acid for wound healing, tissue repair and reducing
 PT hypertrophic scar and keloid formation.
 XX
 PS Claim 3; Page 36-38; 58pp; English.
 XX
 CC cDNA clone 30C includes a coding region for hyaluronan synthase (HAS)
 CC (see AAW26765), an enzyme involved in the synthesis of hyaluronan
 CC (hyaluronic acid) and which has the ability to confer cell adhesion by
 CC the lymphocyte receptor CD44. Clone 30C was isolated using an expression
 CC cloning system developed to isolate cDNA clones that encode proteins that
 CC confer adhesion of the murine T cell lymphoma TK1. A human mesenteric
 CC lymph node expression library was constructed that, upon transfection
 CC into CHO/P cells, yielded clone 30C that mediated rosetting of TK1 cells
 CC to some of the transfectants. The isolated clone can be utilised in a
 CC claimed method for producing HAS in host cells. Such host cells are used
 CC in a claimed method for the production of hyaluronan. Hyaluronan is
 CC useful for wound healing and tissue repair, and can reduce or prevent
 CC hypertrophic scars and keloid formation. It is also used in eye surgery
 CC as a replacement for vitreous fluid
 XX
 SQ Sequence 2116 BP; 330 A; 651 C; 712 G; 423 T; 0 U; 0 Other;
 Query Match 100.0%; Score 20; DB 2; Length 2116;
 Best Local Similarity 100.0%; Pred. No. 3.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGGGCTTGTACAGACTACTT 20
 DB 912 CGGGCTTGTACAGACTACTT 931
 RESULT 7
 ID AAT96713
 XX AAT96713 standard; DNA; 2117 BP.
 XX
 AC AAT96713;
 XX
 XX 22-APR-1998 (first entry)
 XX
 XX Human hyaluronate synthetase coding sequence.
 XX
 KW Hyaluronate synthetase; human; hyaluronic acid; drug preparation;
 KW cosmetic preparation; gene therapy; carcinogenesis; ss.
 XX
 OS Homo sapiens.
 XX

FH Key Location/Qualifiers
 FT CDS 149..1780 /*tag= a
 FT
 XX
 XX WO9738113-A1.
 XX
 XX 16-OCT-1997.
 XX
 XX 31-MAR-1997; 97WO-JP001111.
 XX
 XX 05-APR-1996; 96JP-00084326.
 XX 30-APR-1996; 96JP-00109663.
 XX (SEK) SEIKAGAKU CORP.
 XX
 XX Itano N, Kimata K;
 XX
 XX WPI; 1997-512726/47.
 XX P-PSDB; AAW36503.
 XX
 XX DNA encoding human hyaluronate synthetase - for industrial scale
 PT production of hyaluronic acid used in generating anti-carcinogenic drugs
 PT or for cosmetics.
 XX
 XX Claim 3; Page 23-27; 35pp; Japanese.
 XX
 CC This sequence encodes a human hyaluronate synthetase, and is the coding
 CC sequence of the invention. The encoded enzyme is useful for industrial
 CC scale production of hyaluronic acid for use in the preparation of drugs
 CC and cosmetics. The drugs can also be used in compositions for the
 CC treatment of disorders involving the lowering of hyaluronic acid
 CC production. The peptides may be used for the preparation of antibodies
 CC recognising hyaluronate synthetase, e.g. for diagnostic purposes.
 CC Antisense DNA or RNA corresponding to the DNA can be used in gene therapy
 CC treatment of carcinogenesis
 XX
 SQ Sequence 2117 BP; 306 A; 673 C; 703 G; 435 T; 0 U; 0 Other;
 Query Match 100.0%; Score 20; DB 2; Length 2117;
 Best Local Similarity 100.0%; Pred. No. 3.1;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGGGCTTGTACAGACTACTT 20
 DB 920 CGGGCTTGTACAGACTACTT 939
 RESULT 8
 ID ADL13693/C
 XX ADL13693 standard; DNA; 231222 BP.
 XX
 XX ADL13693;
 XX
 XX 06-MAY-2004 (first entry)
 XX
 XX Osteoarthritis-associated polymorphic nucleotide #225.
 XX
 XX ds; gene; osteopathic; antiinflammatory; antiarthritic; gene therapy;
 KW joint space narrowing; osteophyte development; joint pain;
 KW osteoarthritis; SNP; single nucleotide polymorphism.
 XX
 OS Homo sapiens.
 XX
 XX WO2003054166-A2.
 XX
 XX 03-JUL-2003.
 XX
 XX 19-DEC-2002; 2002WO-US041225.
 XX
 XX 20-DEC-2001; 2001US-0342603P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX

PI Jones KA, Schafer A;
 XX WPI; 2003-559141/52.
 XX
 XX Determining susceptibility of an individual to joint space narrowing,
 PT osteophyte development and/or joint pain comprises identifying whether
 PT the individual has at least one polymorphism in a polynucleotide encoding
 PT a protein.
 XX
 XX Disclosure; SEQ ID NO 225; 297pp; English.
 XX
 XX The invention relates to a method of determining susceptibility of an
 CC individual to joint space narrowing and/or osteophyte development and/or
 CC joint pain comprising identifying whether the individual has at least one
 CC polymorphism in a polynucleotide encoding at least one of the protein
 CC listed in the specification. The methods, composition and agent are
 CC useful for modulating the susceptibility of an individual to joint space
 CC narrowing and/or osteophyte development and/or joint pain that is
 CC associated with a disease, preferably osteoarthritis. The cell line and
 CC the non-human animal are useful for screening for an agent for diagnosing
 CC an individual having susceptibility to joint space narrowing and/or
 CC osteophyte development and/or joint pain. This sequence corresponds to
 CC the polynucleotide encoding a protein listed in the specification. (Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences).
 XX
 XX Sequence 231222 BP; 64006 A; 52087 C; 53478 G; 61650 T; 0 U; 1 Other;

Query Match 100.0%; Score 20; DB 10; Length 231222;
 Best Local Similarity 100.0%; Pred. No. 5.6;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTCTCAGAGCTACTT 20
 |||||
 Db 146240 CGGGCTTGTCTCAGAGCTACTT 146221

RESULT 9
 AAZ10866
 ID AAZ10866 standard; DNA; 30 BP.
 XX
 XX AAZ10866;
 XX
 DT 19-OCT-1999 (first entry)
 XX
 DE PCR primer for the hyaluronate synthase promoter sequence.
 XX
 KW Hyaluronate synthase; promoter; mouse; cell growth inhibitor; PCR primer;
 XX ss.
 XX Synthetic.
 OS Mus sp.
 XX JP11196875-A.
 PN
 XX 27-JUL-1999.
 XX
 PF 14-JAN-1998; 98JP-00006191.
 XX
 PR 14-JAN-1998; 98JP-00006191.
 XX
 PA (SEKG) SEIKAGAKU KOGYO CO LTD.
 XX
 XX WPI; 1999-496653/42.
 DR
 XX New promoter DNA of hyaluronate synthase - used to specifically express
 PT gene with cell growth inhibiting activity.
 PT
 PS Example 1; Page 5; 13pp; Japanese.
 XX
 XX This sequence represents a PCR primer used to isolate the mouse
 CC hyaluronate synthase promoter of the invention. The promoter can be used

CC for specifically expressing a gene having cell growth inhibiting activity
 XX
 XX Sequence 30 BP; 7 A; 8 C; 8 G; 7 T; 0 U; 0 Other;

Query Match 92.0%; Score 18.4; DB 2; Length 30;
 Best Local Similarity 95.0%; Pred. No. 12;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGGCTTGTCTCAGAGCTACTT 20
 |||||
 Db 7 CGGGCTTGTCTCAGAGCTACTT 26

RESULT 10
 ABZ76734
 ID ABZ76734 standard; cDNA; 1752 BP.

XX
 AC ABZ76734;

XX
 DT 01-APR-2000 (first entry)

XX Mouse hyaluronan synthase 1 encoding cDNA SEQ ID NO:1.

XX Mouse; hyaluronan synthase; HAS; HAS1; enzyme; ophthalmological;
 KW osteopathic; antiarthritic; gene therapy; angiogenesis inhibitor;
 KW hyaluronic acid; dihydrazide; bioconjugate; dry eye syndrome; eye;
 KW osteoarthritis; gene; ss.

XX Mus sp.

XX Key Location/Qualifiers
 FH CDS 1..1752
 FT /*tag= a
 FT /product= "hyaluronan synthase 1 (HAS1)"

XX WO2003006068-A1.

XX 23-JAN-2003.

XX 10-JUL-2001; 2001WO-US021785.

XX 10-JUL-2001; 2001WO-US021785.

XX (CLEA-) CLEAR SOLUTIONS BIOTECH INC.

XX Dehazya P, Chen W;

XX WPI; 2003-221664/21.

DR P-PSDB; ABP96028.

XX Novel dihydrazide derivatized hyaluronic acid/nucleic acid bioconjugate
 PT for treating dry eye syndrome, has derivatized hyaluronic acid
 PT crosslinked to nucleic acid encoding protein with hyaluronan synthase
 PT activity.

XX Claim 19; Page 56; 62pp; English.

XX The present invention describes a dihydrazide derivatized hyaluronic acid
 CC (HA)/nucleic acid (NA) bioconjugate (I), with derivatized HA linked to NA
 CC encoding protein with hyaluronan synthase (HAS) activity, where NA has
 CC sequence with 79-85% sequence identity (SI) to a sequence of 1752, 1659
 CC or 1665 nucleotides (see ABZ76734 to ABZ76736) or a sequence encoding
 CC proteins with 66-90% SI to a sequence of 583, 552 or 554 amino acids (see
 CC ABP96028 to ABP96030). (I) has ophthalmological, osteopathic and
 CC antiarthritic activities, and can be used in gene therapy and as an
 CC inhibitor of angiogenesis, and as an inducer of expression of (HA) in
 CC human corneal epithelial cell. (I) is useful for transfected a cell of
 CC an eye with nucleic acid, where the nucleic acid comprises ABZ76734 to
 CC ABZ76736), by contacting the cell with (I) comprising the nucleic acid.
 CC (I) is useful for treating dry eye syndrome in an individual. (I) is
 CC useful in gene therapy applications for the treatment of a variety of
 CC medical conditions including dry eye syndrome or other medical conditions
 CC where an increase in the production of (HA) in the eye would be


```
RESULT 13
AAZ10862
ID AAZ10862 standard; DNA; 2102 BP.
XX
AC AAZ10862;
XX
DT 19-OCT-1999 (first entry)
XX
DE Hyaluronate synthase coding sequence.
XX
KW Hyaluronate synthase; promoter; mouse; cell growth inhibitor; ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 49..1800
FT FT /*tag= a
XX
PN JP11196875-A.
XX
PD 27-JUL-1999.
XX
PF 14-JAN-1998; 98JP-00006191.
XX
PR 14-JAN-1998; 98JP-00006191.
XX
PA (SEBK ) SEIKAGAKU KOGYO CO LTD.
XX
WPI; 1999-496653/42.
XX
P-PSDB; AAY32503.
XX
New promoter DNA of hyaluronate synthase - used to specifically express
gene with cell growth inhibiting activity.
XX
Example 1; Page 8-10; 13pp; Japanese.
XX
This sequence represents the mouse hyaluronate synthase coding sequence,
and was used to isolate the hyaluronate synthase promoter of the
CC invention. The promoter can be used for specifically expressing a gene
CC having cell growth inhibiting activity
XX
SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;
Query Match 92.0%; Score 18.4; DB 2; Length 2102;
Best Local Similarity 95.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 CGGGCTTGTCTCAGAGCTACTT 20
||| ||||| ||||| ||||| |||||
Db 940 CGAGCTTGTCTCAGAGCTACTT 959

RESULT 14
AAZ88199
ID AAZ88199 standard; cDNA to mRNA; 2102 BP.
XX
AC AAZ88199;
XX
DT 27-APR-2000 (first entry)
XX
DE Mouse hyaluronate synthase modified protein HAS1 encoding cDNA.
XX
KW Mouse; murine; hyaluronate synthase; modification; HAS1; HAS2; HAS3;
XX research reagent; biochemical research; medical development; ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 49..1800
FT FT /*tag= a
FT FT /product= "HAS1"
FT FT /note= "modified hyaluronate synthase protein"
XX
PN JP20000116382-A.
XX
PD 25-APR-2000.
XX
PF 13-OCT-1998; 98JP-00291201.
XX
PR 13-OCT-1998; 98JP-00291201.
XX
PA (SEBK ) SEIKAGAKU KOGYO CO LTD.
XX
WPI; 2000-369404/32.
XX
P-PSDB; AAB09948.
XX
DNA used for gene targeting of the hyaluronate synthase gene (HAS1).
```

```
PN JP2000004886-A.
XX
PD 11-JAN-2000.
XX
PF 24-JUN-1998; 98JP-00193788.
XX
PR 24-JUN-1998; 98JP-00193788.
XX
PA (SEBK ) SEIKAGAKU KOGYO CO LTD.
XX
WPI; 2000-140125/13.
XX
P-PSDB; AAY68491.
XX
A hyaluronate synthase modified protein - useful as a research reagent
PT for biochemical research and medical development.
XX
Claim 18; Page 12-14; 30pp; Japanese.
XX
The present invention describes a recombinant protein which consists of
three continuous regions (N-terminal region, internal region and C-
terminal region) where one or two regions among the above three regions
is selected from the three hyaluronate synthase (HAS) modified proteins
CC HAS1, HAS2 and HAS3 (e.g. a protein made from HAS1 amino acids 1 to 71,
CC HAS2 amino acids 52 to 379 and HAS1 amino acids 410 to 583). The protein
CC is useful as a research reagent for biochemical research and medical
CC development. The invention provides a HAS modified protein of increased
CC or lowered activity. The present sequence encodes mouse HAS1
XX
SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;
Query Match 92.0%; Score 18.4; DB 3; Length 2102;
Best Local Similarity 95.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1 CGGGCTTGTCTCAGAGCTACTT 20
||| ||||| ||||| ||||| |||||
Db 940 CGAGCTTGTCTCAGAGCTACTT 959

RESULT 15
AAZ39987
ID AAZ39987 standard; cDNA; 2102 BP.
XX
AC AAZ39987;
XX
DT 19-OCT-2000 (first entry)
XX
DE Murine HAS1 cDNA.
XX
KW Murine; HAS1; hyaluronate synthase; marker gene; intracellular loop;
XX gene targeting; ss.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 49..1800
FT FT /*tag= a
FT FT /product= "HAS1"
XX
PN JP20000116382-A.
XX
PD 25-APR-2000.
XX
PF 13-OCT-1998; 98JP-00291201.
XX
PR 13-OCT-1998; 98JP-00291201.
XX
PA (SEBK ) SEIKAGAKU KOGYO CO LTD.
XX
WPI; 2000-369404/32.
XX
P-PSDB; AAB09948.
XX
DNA used for gene targeting of the hyaluronate synthase gene (HAS1).
```

XX Example; Page 7-9; 14pp; Japanese.
PS
XX
CC This invention describes a novel DNA which contains an heterologous DNA
CC (I) to be introduced to chromosomal DNA of a host cell, a first and
CC second homologous region DNA (II) connected respectively to the 5' and 3'
CC sides of the introduced DNA, and a negative marker gene (III) expressable
CC in the host cell. (I) contains a positive marker gene expressable in the
CC host cell. (I), (II) and the region encoding the intracellular loop of
CC the hyaluronate synthase (HAS) 1 protein from chromosomal DNA is
CC constituted so as to cause a homologous recombination. The DNA can be
CC used for gene targeting of the HAS1 gene. This sequence encodes the
CC murine HAS1 protein described in the method of the invention
XX
SQ Sequence 2102 BP; 399 A; 565 C; 640 G; 498 T; 0 U; 0 Other;
Query Match 92.0%; Score 18.4; DB 3; Length 2102;
Best Local Similarity 95.0%; Pred. No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 940 CGAGCTTGTCTCAGAGCTACTT 959

Search completed: March 13, 2005, 03:46:33
Job time : 17.9978 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:01:16 ; Search time 3.93949 seconds
(without alignments)
8307.054 Million cell updates/sec

Title: US-10-672-399-9

Perfect score: 20

Sequence: 1 cgggcttcagagctactt 20

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	20	100.0	601	4	US-09-949-016-201780 Sequence 201780,
2	20	100.0	2087	4	US-09-949-016-5659 Sequence 5659, Ap
3	20	100.0	2116	3	US-08-635-552A-1 Sequence 1, Appli
4	20	100.0	2117	3	US-09-155-768-1 Sequence 1, Appli
5	20	100.0	14862	4	US-09-949-016-17401 Sequence 17401, A
6	17.4	87.0	91831	4	US-09-949-016-13694 Sequence 13694, A
7	16.4	82.0	3831	4	US-09-360-394C-1 Sequence 1, Appli
8	15.8	79.0	400	4	US-08-956-171E-4058 Sequence 4058, Ap
9	15.8	79.0	400	4	US-08-781-986A-4058 Sequence 4058, Ap
10	15.8	79.0	9623	4	US-08-956-171E-166 Sequence 166, App
11	15.8	79.0	9623	4	US-08-781-986A-166 Sequence 166, App
12	15.8	79.0	193169	4	US-09-949-016-15091 Sequence 20, Appl
13	15.8	78.0	246240	2	US-08-724-394A-21 Sequence 21, Appl
14	15.8	79.0	246240	2	US-08-724-394A-21 Sequence 22, Appl
15	15.8	79.0	246240	2	US-08-724-394A-22 Sequence 77, Appl
16	15.4	77.0	1642	4	US-10-000-489-77 Sequence 12802, A
17	15.4	77.0	13782	4	US-09-949-016-12802 Sequence 136494,
18	15.2	76.0	601	4	US-09-949-016-136494 Sequence 1506, Ap
19	15.2	76.0	750	4	US-09-489-039A-1506 Sequence 5123, Ap
20	15.2	76.0	1844	4	US-09-949-016-5123 Sequence 1, Appli
21	15.2	76.0	1856	1	US-08-706-214-1 Sequence 475, App
22	15.2	76.0	1856	4	US-09-706-214-1 Sequence 1, Appli
23	15.2	76.0	2492	3	US-09-327-984A-1 Sequence 12217, A
24	15.2	76.0	17472	4	US-09-949-016-12217 Sequence 16865, A
25	15.2	76.0	17476	4	US-09-949-016-16865 Sequence 12585, A
26	15.2	76.0	19854	4	US-09-949-016-12585 Sequence 17306, A
27	15.2	76.0	19854	4	US-09-949-016-17306

28	15.2	76.0	21721	4	US-09-269-939A-41 Sequence 41, Appl
29	15.2	76.0	22976	4	US-09-269-939A-19 Sequence 19, Appl
30	15.2	76.0	23187	4	US-09-499-522-1 Sequence 1, Appli
31	15.2	76.0	97423	4	US-09-949-016-12742 Sequence 12742, A
32	15.2	76.0	97424	4	US-09-949-016-15576 Sequence 15576, A
33	15.2	76.0	108440	4	US-09-949-016-12065 Sequence 12065, A
34	15.2	76.0	108441	4	US-09-949-016-14090 Sequence 14090, A
C 35	15.2	76.0	1830121	4	US-09-557-884-1 Sequence 1, Appli
C 36	15.2	76.0	1830121	4	US-09-643-990A-1 Sequence 1, Appli
C 37	15	75.0	639	3	US-09-171-517B-1 Sequence 771, App
38	15	75.0	5351	3	US-09-221-017B-771 Sequence 82354, A
39	14.8	74.0	601	4	US-09-949-016-82354 Sequence 82355, A
40	14.8	74.0	601	4	US-09-949-016-82355 Sequence 82356, A
41	14.8	74.0	601	4	US-09-949-016-82356 Sequence 82357, A
42	14.8	74.0	601	4	US-09-949-016-82357 Sequence 82358, A
43	14.8	74.0	601	4	US-09-949-016-82358 Sequence 82359, A
44	14.8	74.0	601	4	US-09-949-016-82359 Sequence 85681, A
45	14.8	74.0	601	4	US-09-949-016-85681

ALIGNMENTS

RESULT 1
US-09-949-016-201780/c
; Sequence 201780, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 201780
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-201780

Query Match 100.0% Score 20; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTCTCAGAGCTACTT 20
|||||
DB 222 CGGGCTTCTCAGAGCTACTT 203
|||||

RESULT 2
US-09-949-016-5659
; Sequence 5659, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5659
; LENGTH: 2087
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5659

Query Match      100.0%; Score 20; DB 4; Length 2087;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTGACAGCTACTT 20
Db 912 CGGGCTTGTGACAGCTACTT 931

RESULT 3
US-08-635-552A-1
; Sequence 1, Application US/08635552A
; Patent No. 6423514
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; TITLE OF INVENTION: Mammalian Hyaluronan Synthases, Nucleic
; TITLE OF INVENTION: Acids and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/635,552A
; FILING DATE: 22-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Brook, David E.
; REGISTRATION NUMBER: 22,592
; REFERENCE/DOCKET NUMBER: LKS95-07
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2116 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36...1769
US-08-635-552A-1

Query Match      100.0%; Score 20; DB 3; Length 2116;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTGACAGCTACTT 20
Db 912 CGGGCTTGTGACAGCTACTT 931

RESULT 4
US-09-155-768-1
; Sequence 1, Application US/09155768A
; Patent No. 6162908
; GENERAL INFORMATION:
; APPLICANT: SEIKAGAKU CORPORATION
; TITLE OF INVENTION: POLYPEPTIDE OF HYALURONAN SYNTHASE OF
; TITLE OF INVENTION: HUMAN-ORIGIN AND DNA ENCODING THE SAME
; FILE REFERENCE: TOYAM37.001APC
; CURRENT APPLICATION NUMBER: US/09/155,768A
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: JP 8-084326
; EARLIER FILING DATE: 1996-04-05
; EARLIER APPLICATION NUMBER: JP 8-109663
; EARLIER FILING DATE: 1996-04-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2117
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (149)...(1777)
US-09-155-768-1

Query Match      100.0%; Score 20; DB 3; Length 2117;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTGACAGCTACTT 20
Db 920 CGGGCTTGTGACAGCTACTT 939

RESULT 5
US-09-949-016-17401
; Sequence 17401, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17401
; LENGTH: 14862
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17401

Query Match      100.0%; Score 20; DB 4; Length 14862;
Best Local Similarity 100.0%; Pred. No. 0.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTGACAGCTACTT 20
Db 8955 CGGGCTTGTGACAGCTACTT 8974

RESULT 6
US-09-949-016-13694
; Sequence 13694, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; PRIOR FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13694
 ; LENGTH: 91831
 ; TYPE: DNA
 ; ORGANISM: Human
 ; ORGANISM: Human
 US-09-949-016-13694

Query Match 87.0%; Score 17.4; DB 4; Length 91831;
 Best Local Similarity 94.7%; Pred. No. 26;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGCTTGTTCAGAGCTACTT 20
 DB 55477 GGGCTTGTTCAGAGCTACTT 55495

RESULT 7
 US-09-360-394C-1
 ; Sequence 1, Application US/09360394C
 ; Patent No. 6610509
 ; GENERAL INFORMATION:
 ; APPLICANT: Sandell, Linda
 ; APPLICANT: Xie, Wei-Pen
 ; TITLE OF INVENTION: METHODS OF TARGETED EXPRESSION BY THE CD-RAP GENE PROMOTER
 ; FILE REFERENCE: BJCH 10026
 ; CURRENT APPLICATION NUMBER: US/09/360,394C
 ; CURRENT FILING DATE: 1999-07-23
 ; PRIOR APPLICATION NUMBER: 60-094,092
 ; PRIOR FILING DATE: 1998-07-24
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 3831
 ; TYPE: DNA
 ; ORGANISM: Mus musculus;
 ; FEATURE:
 ; NAME/KEY: Unsure
 ; LOCATION: (3598)..(3598)
 ; OTHER INFORMATION: n=unknown
 ; FEATURE:
 ; NAME/KEY: Unsure
 ; LOCATION: (3591)..(3591)
 ; OTHER INFORMATION: n=unknown
 ; FEATURE:
 ; NAME/KEY: Unsure
 ; LOCATION: (3624)..(3624)
 ; OTHER INFORMATION: n=unknown
 ; FEATURE:
 ; NAME/KEY: Unsure
 ; LOCATION: (3790)..(3790)
 ; OTHER INFORMATION: n=unknown
 ; FEATURE:
 ; NAME/KEY: Unsure
 ; LOCATION: (3805)..(3805)
 ; OTHER INFORMATION: n=unknown
 ; FEATURE:
 ; NAME/KEY: Unsure
 ; LOCATION: (3821)..(3821)
 ; OTHER INFORMATION: n=unknown
 US-09-360-394C-1

Query Match 82.0%; Score 16.4; DB 4; Length 3831;
 Best Local Similarity 94.4%; Pred. No. 50;
 Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGCTTGTTCAGAGCTACT 19
 DB 1176 GGGCTTGTTCAGAGCTACT 1193

RESULT 8
 US-08-956-171E-4058
 ; Sequence 4058, Application US/08956171E
 ; Patent No. 6593114
 ; GENERAL INFORMATION:
 ; APPLICANT: Charles Kunech
 ; APPLICANT: Gil H. Choi
 ; APPLICANT: Patrick S. Dillon
 ; APPLICANT: Craig A. Rosen
 ; APPLICANT: Steven C. Barash
 ; APPLICANT: Michael R. Fannon

; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
 ; NUMBER OF SEQUENCES: 5256
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Human Genome Sciences, Inc.
 ; STREET: 9410 Key West Avenue
 ; CITY: Rockville
 ; STATE: Maryland
 ; COUNTRY: USA
 ; ZIP: 20850

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
 ; COMPUTER: HP Vectra 486/33
 ; OPERATING SYSTEM: MSDOS version 6.2
 ; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/956,171E
 ; FILING DATE: 20-Oct-1997
 ; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/009,861
 ; FILING DATE: January 5, 1996
 ; APPLICATION NUMBER: 08/781,986
 ; FILING DATE: January 3, 1997

; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mark J. Hyman
 ; REGISTRATION NUMBER: 46,789

; REFERENCE/DOCKET NUMBER: PB248PI
 ; TELEPHONE: (240) 314-1224
 ; TELEFAX: (301) 309-8439

; INFORMATION FOR SEQ ID NO: 4058:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 400 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear

; SEQUENCE DESCRIPTION: SEQ ID NO: 4058:
 US-08-956-171E-4058

Query Match 79.0%; Score 15.8; DB 4; Length 400;
 Best Local Similarity 89.5%; Pred. No. 68;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGCTTGTTCAGAGCTACT 19
 DB 66 GGGCTTGTTCAGAGCTACT 84

RESULT 9
 US-08-781-986A-4058
 ; Sequence 4058, Application US/08781986A
 ; Patent No. 6737248
 ; GENERAL INFORMATION:

```
;
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781.986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 4058:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-781-986A-4058
;
; Query Match 79.0%; Score 15.8; DB 4; Length 400;
; Best Local Similarity 89.5%; Pred. No. 68;
; Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; Qy 1 CGGGCTTGTCTCAGCGTACT 19
; Db 66 CGGGCTTGTCTCAGCGTACT 84
;
; RESULT 10
; US-08-956-171E-166
; Sequence 166, Application US/08956171E
; Patent No. 6593114
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; Gil H. Choi
; Patrick S. Dillon
; Craig A. Rosen
; Steven C. Barash
; Michael R. Fannon
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5256
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 166:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9623 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-781-986A-166
```

```
;
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 166:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9623 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 166:
;
; US-08-956-171E-166
;
; Query Match 79.0%; Score 15.8; DB 4; Length 9623;
; Best Local Similarity 89.5%; Pred. No. 1.3e+02;
; Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
;
; Qy 1 CGGGCTTGTCTCAGCGTACT 19
; Db 7504 CGGGCTTGTCTCAGCGTACT 7522
;
; RESULT 11
; US-08-781-986A-166
; Sequence 166, Application US/08781986A
; Patent No. 6737248
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 166:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9623 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
; US-08-781-986A-166
```



```

* OPERATING SYSTEM: PC-DOS/MS-DOS
* SOFTWARE: PatentIn Release #1.0, Version #1.30
* CURRENT APPLICATION DATA:
* APPLICATION NUMBER: US/08/724,394A
* FILING DATE: 01-OCT-1996
* CLASSIFICATION: 536
* ATTORNEY/AGENT INFORMATION:
* NAME: Fitts, Renee A.
* REGISTRATION NUMBER: 35,136
* REFERENCE/DOCKET NUMBER: 017957-000100
* TELECOMMUNICATION INFORMATION:
* TELEPHONE: 415-576-0200
* TELEFAX: 415-576-0300
* INFORMATION FOR SEQ ID NO: 20:
* SEQUENCE CHARACTERISTICS:
* LENGTH: 246240 base pairs
* TYPE: nucleic acid
* STRANDEDNESS: not relevant
* TOPOLOGY: not relevant
* MOLECULE TYPE: cdna
* FEATURE:
* NAME/KEY: misc feature
* LOCATION: 1..246240
* OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-20

Query Match 79.0%; Score 15.8; DB 2; Length 246240;
Best Local Similarity 89.5%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0

Qy 2 GGGCTTGTCAGAGCTACTT 20
Db 60431 GGGCTTGTCAGAGGTATT 60449

RESULT 14
US-08-724-394A-21
; Sequence 21, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Kronmal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Tsuchihashi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300

```

; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-21

Query Match 79.0%; Score 15.8; DB 2; Length 246240;
Best Local Similarity 89.5%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCTTGTGCAGAGCTACTT 20
|||||
Db 60431 GGGCTTGTGCAGAGGTATT 60449

RESULT 15
US-08-724-394A-22
; Sequence 22, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
; APPLICANT: Feder, John N.
; APPLICANT: Krommal, Gregory S.
; APPLICANT: Lauer, Peter M.
; APPLICANT: Ruddy, David A.
; APPLICANT: Thomas, Winston
; APPLICANT: Teuchihaahi, Zenta
; APPLICANT: Wolff, Roger K.
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
; TITLE OF INVENTION: Sequences and Antibodies Thereto
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,394A
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Pitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 017957-000100
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 246240 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..246240
; OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-22

Query Match 79.0%; Score 15.8; DB 2; Length 246240;
Best Local Similarity 89.5%; Pred. No. 2.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 GGGCTTGTGCAGAGCTACTT 20
|||||
Db 60431 GGGCTTGTGCAGAGGTATT 60449

Search completed: March 13, 2005, 08:12:26
Job time : 7.93949 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:08:52 ; Search time 82.4452 Seconds
(without alignments)
9233.835 Million cell updates/sec

Title: US-10-672-399-9

Perfect score: 20

Sequence: 1 cgggcttcagagctactt 20

Scoring table: IDENTITY_NUC

Gapop 10_0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	551	4	BG707429
2	20	100.0	551	4	BG707608
3	20	100.0	610	5	BX327795
4	20	100.0	739	9	AY421570
5	20	100.0	782	4	BI753116
6	20	100.0	896	5	BX371636
7	20	100.0	1084	4	BM544718
8	20	100.0	1728	9	AY421569
9	20	100.0	2072	3	CR602106
10	18.4	92.0	166	5	BM936724
11	18.4	92.0	495	7	CN664556
12	18.4	92.0	1746	9	AY421571
13	18.4	92.0	3542	3	AK053726
14	17.4	87.0	442	5	BX264217
15	17.4	87.0	500	7	CK675190
16	17.4	87.0	502	5	BX372063
17	17.4	87.0	506	5	BQ169305
18	17.4	87.0	514	1	AL927904
19	17.4	87.0	570	1	AI878069
20	17.4	87.0	644	7	CF253438
21	17.4	87.0	681	6	CD777815
22	17.4	87.0	686	5	BU301603
23	17.4	87.0	722	5	BU306034
24	17.4	87.0	723	5	BU302845

c	25	17.4	87.0	743	5	BU269373
	26	17.4	87.0	754	6	CD777811
	27	17.4	87.0	819	7	CK687847
c	28	17.4	87.0	855	1	AJ444643
	29	17.4	87.0	868	7	CO924080
	30	17.4	87.0	888	7	CK017913
c	31	17.4	87.0	891	5	BU230362
	32	17.4	87.0	1018	5	BQ891951
c	33	17.4	87.0	336	7	CK116412
	34	17.4	87.0	456	7	CK116743
c	35	17.4	87.0	514	1	AI162566
	36	17.4	87.0	721	5	BU884547
	37	16.8	84.0	356	2	BF708639
c	38	16.8	84.0	437	9	FR002608
	39	16.8	84.0	619	9	CL588198
c	40	16.8	84.0	622	2	BE289996
	41	16.8	84.0	736	9	AG411055
c	42	16.8	84.0	770	9	CL516805
	43	16.8	84.0	817	5	BU451953
c	44	16.8	84.0	1036	9	CNS04RFO
	45	16.4	82.0	303	1	AA781062

ALIGNMENTS

RESULT 1
BG707429 551 bp mRNA linear EST 07-MAY-2001
LOCUS 602672848F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4795643 5',
DEFINITION mRNA sequence.
ACCESSION BG707429
VERSION BG707429.1 GI:13983769
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 551)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAW10678 row: p column: 12
High quality sequence stop: 551.
Location/Qualifiers
1. 551
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4795643"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH MGC 96"
/note="Organ: brain; Vector: pbluescriptR (modified
pbluescript KS+); Site: 1: BamHI; Site 2: SalI-XhoI
(gtcgag); Oligo-dr primed using primer
5'-TTTTTTTTTTTTTTN-3', size-selected for average
insert size 2.3 kb and normalized to 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

FEATURES

source

```

ORIGIN
Query Match      100.0%; Score 20; DB 4; Length 551;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTCTCAGAGCTACTT 20
    |||||
DB 248 CGGGCTTGTCTCAGAGCTACTT 267

RESULT 2
LOCUS BG707608 551 bp mRNA linear EST 07-MAY-2001
DEFINITION 602670672F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4793704 5',
            mRNA sequence.
ACCESSION BG707608
VERSION BG707608.1 GI:13984125
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 551)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 2, 2003 this sequence version replaced gi:30340686.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
9502.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?e=CS0BAG028ZH06_CS03640_1&c=9502.r

FEATURES
            source
            Location/Qualifiers
                1..610
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                /tissue_type="PLACENTA COT 25-NORMALIZED"
                /note="1st strand cDNA was primed with a NotI-oligo (dT)
                primer. Five prime end enriched, double-strand cDNA was
                digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match      100.0%; Score 20; DB 5; Length 610;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTCTCAGAGCTACTT 20
    |||||
DB 219 CGGGCTTGTCTCAGAGCTACTT 238

RESULT 4
LOCUS AY421570 739 bp DNA linear GSS 17-DEC-2003
DEFINITION Pan troglodytes HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION AY421570
VERSION AY421570.1 GI:39748429
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 739)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 739)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

ORIGIN
Query Match      100.0%; Score 20; DB 4; Length 551;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTCTCAGAGCTACTT 20
    |||||
DB 248 CGGGCTTGTCTCAGAGCTACTT 267

RESULT 3
LOCUS BX327795 610 bp mRNA linear EST 08-APR-2004
DEFINITION BX327795 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

```

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sniensky,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.

FEATURES
 source
 1..739
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..739
 /genes="HAS1"
 /locus_tag="HCM7601"

ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 739;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTACAGCTACTT 20
 |||||
 Db 132 CGGGCTTGTACAGCTACTT 151

RESULT 5

BI753116 782 bp mRNA linear EST 25-SEP-2001
 LOCUS 603025961F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5196399 5',
 DEFINITION mRNA sequence.

ACCESSION BI753116
 VERSION BI753116.1 GI:15744694
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 782)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Life Technologies, Inc.

CDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM1491 row: j column: 16

High quality sequence stop: 780.

Location/Qualifiers

FEATURES

source
 1..782
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5196399"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_114"
 /note="Organ: brain; Vector: pCMV-Sport6; Site 1: NotI;
 Site 2: EcoRV (destroyed); RNA source anonymous pool of 6
 male brains, age range 23-27 yo. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.5 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 019. Note:
 this is a NIH_MGC Library."

ORIGIN

Query Match 100.0%; Score 20; DB 4; Length 782;

Best Local Similarity 100.0%; Pred. No. 12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CGGGCTTGTACAGCTACTT 20
 |||||
 Db 212 CGGGCTTGTACAGCTACTT 231

RESULT 6

BI7531636/c 896 bp mRNA linear EST 27-APR-2004
 LOCUS BI7531636 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 DEFINITION clone CSODI027YJ02 3-PRIME, mRNA sequence.

ACCESSION BI7531636
 VERSION BI7531636.2 GI:46616231
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 896)
 Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 8, 2003 this sequence version replaced gi:30450029.

CONTACT: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

9502.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?g=CS0BAI019ZC12_CS01819_1&c=9502.r

FEATURES

source
 Location/Qualifiers
 1..896
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODI027YJ02"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 896;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTACAGCTACTT 20
 |||||
 Db 356 CGGGCTTGTACAGCTACTT 337

RESULT 7

BM544718 1084 bp mRNA linear EST 20-FEB-2002
 LOCUS BM544718
 DEFINITION AGENCOCURT_6494603 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5727446
 5', mRNA sequence.

ACCESSION BM544718
 VERSION BM544718.1 GI:18776197
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Invitrogen
cDNA library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: Agencourt Bioscience Corporation
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12721 row: i column: 15
High quality sequence start: 5
High quality sequence stop: 689.
FEATURES
source
1. .1084
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5727446"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH MGC 124"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: EcoRV
(destroyed); Site_2: NotI; RNA source male hippocampus,
age 27. Library is oligo-dT primed and directionally
cloned (EcoRV site is destroyed upon cloning). Average
insert size 1.4 kb, insert size range 0.9-4 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 012."
ORIGIN
Query Match 100.0%; Score 20; DB 4; Length 1084;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTCTCAGAGCTACTT 20
|||||
Db 276 CGGGCTTGTCTCAGAGCTACTT 295

RESULT 8
AY421569
LOCUS
DEFINITION Homo sapiens HAS1 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY421569
VERSION AY421569.1 GI:39748428
KEYWORDS GSS.
SOURCE AY421569.1
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini, Homiidae; Homo.
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 1728)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1. .1728
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1728
/feature="HAS1"
/locus_tag="HCM7601"
ORIGIN
Query Match 100.0%; Score 20; DB 9; Length 1728;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTCTCAGAGCTACTT 20
|||||
Db 868 CGGGCTTGTCTCAGAGCTACTT 887

RESULT 9
CR602106
LOCUS
DEFINITION full-length cDNA clone CS0DI027YJ02 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR602106
VERSION CR602106.1 GI:50482913
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 2072)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1. .2072
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI027YJ02"
/tissue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 100.0%; Score 20; DB 3; Length 2072;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGGCTTGTCTCAGAGCTACTT 20
|||||
Db 925 CGGGCTTGTCTCAGAGCTACTT 944

RESULT 10
BM936724
LOCUS

```

DEFINITION UI-M-BH3-ar-d-04-0-UI.r1 NIH_BMAP_M_S4 Mus musculus cDNA clone
 UI-M-BH3-ar-d-04-0-UI.5', mRNA sequence.
 ACCESSION BM936724
 VERSION BM936724.1 GI:19395876
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 166)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene
 discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 9704477
 PUBMED 8889548
 COMMENT Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mEST@mail.nih.gov
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 REVERSE.

FEATURES

Location/Qualifiers

1..166
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH3-ar-d-04-0-UI"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NIH_BMAP_M_S4"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; The
 NIH_BMAP_M_S4 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged
 normalized libraries from ten regions of the mouse brain
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,
 cortex, amygdala, basal ganglia, pineal gland, striatum,
 hippocampus) after a series of subtractions to reduce the
 representation of cDNAs from which ESTs had already been
 generated. The following serially subtracted libraries
 were generated in this process: NIH_BMAP_M_S4,
 NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
 NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
 (NIH_BMAP_M_S4) was constructed as follows: PCR amplified
 cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
 NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
 was used as a driver in a hybridization with a pool of
 the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
 libraries in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library)
 was purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 NIH_BMAP_M_S4 library. This procedure has been previously
 described (Bonaldo, Lennon and Soares, Genome Research
 6:791-806, 1996)."

source

Db

RESULT 11

CN664556/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1..495
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="NIA:A0814F02 IMAGE:30758557"
 /tissue_type="whole embryo including extraembryonic
 tissues at 13.5-days postcoitum"
 /dev_stage="E13.5"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse E13.5 whole embryo cDNA library
 (long)"
 /note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
 Site 2: NotI; Mouse cDNA project by the Laboratory of
 Genetics, National Institute on Aging (NIA), Intramural
 Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
 This is a long-transcript enriched cDNA library (Ref.
 Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
 RNAs were extracted from 1 embryo at 13.5-days postcoitum.
 Double-stranded cDNAs were synthesized with an Oligo(dT)
 primer [Invitrogen]:
 5'-pGACTAGTCTAGATCGGAGCGGCCCTTTT-3' from
 3ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to Lona-linker LK-Sal4, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Taq polymerase (Takara) with a primer Sal4-S. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
 vector. The DH10B E. coli host was transformed with the
 ligation mixture by the standard chemical method. The
 average insert size is about 3.0Kb. The library was
 constructed by Yulan Piao."

ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 166;
 Best Local Similarity 95.0%; Pred. No. 66;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGCTTGTCTAGACTT 20

|||||

ORIGIN

91 CGAGCTTGTCTAGACTT 110

CN664556

A0814F02-5 NIA Mouse E13.5 whole embryo cDNA library (long) Mus

musculus cDNA clone NIA:A0814F02 IMAGE:30758557 5', mRNA sequence.

CN664556

CN664556.1 GI:47431007

EST.

Mus musculus (house mouse)

Mus musculus

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Sharov,A.A., Piao,Y., Matoba,R., Dudekula,D.B., Qian,Y.,

VanBuren,V., Falco,G., Martin,P.R., Stagg,C.A., Bassey,U.C.,

Wang,Y., Carter,M.G., Hamatani,T., Aiba,K., Akutsu,H., Sharova,L.,

Tanaka,T.S., Kimber,W.L., Yoshikawa,T., Jaradat,S.A., Pantano,S.,

Nagaraja,R., Boheler,K.R., Taub,D., Hodes,R.J., Longo,D.L.,

Schlessinger,D., Keller,J., Klotz,E., Kelsoe,G., Umezawa,A.,

Vescovi,A.L., Rosant,J., Kunath,T., Hogan,B.L., Curci,A.,

D'Urso,M., Kelsoe,J., Hide,W. and Ko,M.S.

Transcriptome analysis of mouse stem cells and early embryos

PLOS Biol. 1 (3), 410-419 (2003)

Contact: Dawood B. Dudekula

Laboratory of Genetics

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@lgsun.grc.nia.nih.gov

Plate: A0814 row: F column: 02

Seq primer: M13 Reverse

High quality sequence stop: 495

POLYA=No.

[illegible]

/db_xref="taxon:10090"
/clone="E130302P16"
/tissue_type="eyeball"
BX264217 AGENAE full-length enriched mouse cDNA library"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
61..456

CDS

/note="unnamed protein product; hyaluronan synthase1
(MGD|MG1:106590, GB|NM_008215; evidence: BLASTN, 99%,
match=762)
putative"
/codon_start=1
/protein_id="BAC35492.1"
/db_xref="GI:26343671"
/translation="MRQMPKPEARCCGILARRALTIIFALLILGLMTWAYAGVP
LASDRYLAFGLYGAFLSAHLVAOSLFAYLEHRRVAAARRSLAKGPLDAATARSVA
LTISAYQBDPAVLRLQCLTSARALLYPHTR"

ORIGIN

Query Match 92.0%; Score 18.4; DB 3; Length 3542;
Best Local Similarity 95.0%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGGCTTGTCCAGAGCTACTT 20

DB 2519 CGAGCTTGTCCAGAGCTACTT 2538

RESULT 14

BX264217/c
LOCUS
DEFINITION
BX264217 AGENAE Gallus gallus multi-tissues normalized and
once-subtracted cDNA library (gal); Gallus gallus cDNA clone
gcal0014c.h.13 5prim, mRNA sequence.

ACCESSION

VERSION
BX264217.1 GI:28586815

KEYWORDS

SOURCE

ORGANISM

Gallus gallus (chicken)
Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Construction and primary characterization of chicken normalized
multi-tissue cDNA libraries
Unpublished (2003)
Contact: Douaire M
INRA, UMR INRA-ENSAR Genetique Animale
65, rue de Saint-Brieuc, RENNES cedex, 35042, FRANCE
Tel: +33 (0) 2.23.48.54.63
Fax: +33 (0) 2.23.48.54.70
Email: Madeleine.Douaire@roazhon.inra.fr
Sequence cleaned of vector, adaptor and repetitions. Contact us
at signasupport@jouy.inra.fr to obtain the chromatogram of this
sequence.

Seq primer: M13R. h column: 13

FEATURES

source

1..442
Location/Qualifiers
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="gcal0014c.h.13"
/tissue_type="multi-tissues"
/dev_stage="from embryos to adults"
/lab_host="DH10B"
/clone_lib="AGENAE Gallus gallus multi-tissues normalized
and once-subtracted cDNA library (gal)"
/note="vector: pTYR3-pac; tissues: adipose tissue, brain,
kidney, liver, multi-tissues, muscle, ovary, testis, bone
marrow, caecum, duodenum, embryos, fabricius gland,
granulosa, hypothalamus, ileon, jejunum, oviduct,

pancreas, skin, spleen, thymus, utero-vaginal gland,
pituitary gland, hematopoietic progenitor cells, small
follicle. Clone distribution : AGENAE Resource centre.
Francois PUMI, Francois.Piumi.inra.fr, INRA, CEA
Radiobiologie et Etude du genome (LREG), Domaine de
Vilvert, 78352, Jouy-en-Josas cedex, FRANCE"

ORIGIN

Query Match 87.0%; Score 17.4; DB 5; Length 442;
Best Local Similarity 94.7%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGGCTTGTCCAGAGCTACT 19

DB 216 CGGGCTTGTCCAGTCTACT 198

RESULT 15

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

construct full length enriched cDNA library and also served as template to synthesize complex first strand cDNA probe. Two high density colony arrays were made from over 110K cDNA clones and hybridized with the probes. Low intensity clones were selected as they represented rare expressed clones. The hybridization intensities for all clones span from 0 to 1.8 million counts and the low abundant class ranged from 0 to 13,000."

ORIGIN

Query Match 87.0%; Score 17.4; DB 7; Length 500;
 Best Local Similarity 94.7%; Pred. No. 2.9e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCGGCTTGTTCAGAGCTACT 19
 |||||
 Db 480 CCGGCTTGTTCAGAGCTTCT 498

Search completed: March 13, 2005, 11:55:05
 Job time : 87.4452 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 08:01:42 ; Search time 14.0784 Seconds
(without alignments)
8450.560 Million cell updates/sec

Title: US-10-672-399-9

Perfect score: 20

Sequence: 1 cgggcttgctcagactactt 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 5537552 seqs, 2974263231 residues

Total number of hits satisfying chosen parameters: 11075104

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:**

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US05_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	18	US-10-672-399-9
2	20	100.0	490	10	US-09-918-995-27210
3	20	100.0	1065	18	US-10-672-399-7
4	20	100.0	1071	18	US-10-672-399-3
5	20	100.0	1083	18	US-10-672-399-5
6	20	100.0	1073	18	US-10-672-399-1
7	20	100.0	2116	13	US-10-042-523-1
8	18.4	92.0	1752	10	US-09-902-939-1
9	16.4	82.0	575	18	US-10-425-115-53796
10	16.4	82.0	2928	18	US-10-437-963-33800
11	16.4	82.0	21410	13	US-10-087-192-1231

c	12	16.4	82.0	30486	17	US-10-085-117-19	Sequence 19, Appl
	13	15.8	79.0	201	19	US-10-741-600-37582	Sequence 37582, A
	14	15.8	79.0	213	18	US-10-425-115-167738	Sequence 167738, A
	15	15.8	79.0	316	11	US-09-732-627A-2660	Sequence 2660, Ap
	16	15.8	79.0	400	8	US-08-781-986A-4058	Sequence 4058, Ap
	17	15.8	79.0	400	17	US-10-329-624-4058	Sequence 4058, Ap
	18	15.8	79.0	477	9	US-09-864-761-1728	Sequence 1728, A
	19	15.8	79.0	582	9	US-09-864-761-12208	Sequence 12208, A
	20	15.8	79.0	913	13	US-10-027-632-161814	Sequence 161814, A
	21	15.8	79.0	913	17	US-10-027-632-161814	Sequence 161814, A
	22	15.8	79.0	1607	17	US-10-425-114-13595	Sequence 13595, A
c	23	15.8	79.0	1953	17	US-10-282-122A-7912	Sequence 7912, Ap
	24	15.8	79.0	3017	18	US-10-425-115-104761	Sequence 104761, A
	25	15.8	79.0	9623	8	US-08-781-986A-166	Sequence 166, App
	26	15.8	79.0	9623	17	US-10-329-624-166	Sequence 166, App
c	27	15.8	79.0	43975	19	US-10-741-600-17669	Sequence 17669, A
	28	15.8	79.0	50720	18	US-10-322-281-621	Sequence 621, App
	29	15.8	79.0	77834	13	US-10-087-192-343	Sequence 343, App
c	30	15.8	79.0	115863	18	US-10-723-860-2504	Sequence 2504, Ap
c	31	15.8	79.0	235033	15	US-10-301-844-1	Sequence 1, Appli
c	32	15.8	79.0	237326	15	US-10-301-844-2	Sequence 2, Appli
c	33	15.4	77.0	598	13	US-10-027-632-213175	Sequence 213175, A
	34	15.4	77.0	598	17	US-10-027-632-213175	Sequence 213175, A
	35	15.4	77.0	1211	17	US-10-260-238-308	Sequence 308, App
	36	15.4	77.0	1642	10	US-09-992-600A-77	Sequence 77, Appl
	37	15.4	77.0	1642	10	US-09-924-340-77	Sequence 77, Appl
	38	15.4	77.0	1642	10	US-09-992-095B-77	Sequence 77, Appl
	39	15.4	77.0	1642	10	US-09-999-570-77	Sequence 77, Appl
	40	15.4	77.0	1642	14	US-10-000-489-77	Sequence 77, Appl
	41	15.4	77.0	1642	14	US-10-000-986-77	Sequence 77, Appl
	42	15.4	77.0	1642	16	US-10-154-678-77	Sequence 77, Appl
	43	15.4	77.0	1642	16	US-10-001-142-77	Sequence 77, Appl
	44	15.4	77.0	1642	19	US-10-838-854-77	Sequence 77, Appl
	45	15.4	77.0	1655	18	US-10-437-963-97018	Sequence 97018, A

ALIGNMENTS

RESULT 1
US-10-672-399-9
; Sequence 9, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-9

Query Match 100.0%; Score 20; DB 18; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGGCTTGTCAGAGCTACTT 20
Db 1 CCGGCTTGTCAGAGCTACTT 20

RESULT 2
US-09-918-995-27210
; Sequence 27210, Application US/09918995
; Publication No. US200300073623A1
; GENERAL INFORMATION:

```
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27210
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-27210

Query Match      100.0%; Score 20; DB 10; Length 490;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 335 CGGGCTTGTCTCAGAGCTACTT 354

RESULT 3
US-10-672-399-7
; Sequence 7, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 7
; LENGTH: 1065
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-7

Query Match      100.0%; Score 20; DB 18; Length 1065;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 877 CGGGCTTGTCTCAGAGCTACTT 896

RESULT 4
US-10-672-399-3
; Sequence 3, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
```

```
; LENGTH: 1071
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-3

Query Match      100.0%; Score 20; DB 18; Length 1071;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 877 CGGGCTTGTCTCAGAGCTACTT 896

RESULT 5
US-10-672-399-5
; Sequence 5, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5
; LENGTH: 1083
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-5

Query Match      100.0%; Score 20; DB 18; Length 1083;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 877 CGGGCTTGTCTCAGAGCTACTT 896

RESULT 6
US-10-672-399-1
; Sequence 1, Application US/10672399
; Publication No. US20050003368A1
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-1

Query Match      100.0%; Score 20; DB 18; Length 1737;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 877 CGGGCTTGTCTCAGAGCTACTT 896

RESULT 7
```

```

US-09-902-939-1

Query Match          92.0%; Score 18.4; DB 10; Length 1752;
Best Local Similarity 95.0%; Pred. No. 11;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CGGGCTTGTGCAGAGCTACTT 20
   ||| ||||| ||||| |||||
DB 892 CGAGCTTGTGCAGAGCTACTT 911

RESULT 9
US-10-425-115-53796
; Sequence 53796, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 53796
; LENGTH: 575
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_149063C.1
US-10-425-115-53796

Query Match          82.0%; Score 16.4; DB 18; Length 575;
Best Local Similarity 94.4%; Pred. No. 1.2e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 GCCTTGTGCAGAGCTACTT 20
   ||| ||||| ||||| |||||
DB 480 GCCTTGCAGAGCTACTT 497

RESULT 10
US-10-437-963-33800/c
; Sequence 33800, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 33800
; LENGTH: 2928
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37878C.1
US-10-437-963-33800

Query Match          82.0%; Score 16.4; DB 18; Length 2928;
Best Local Similarity 94.4%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 2 GGGCTTGTCTCAGAGCTACT 19
|||||
Db 2706 GGGCTTGTCTCAGAGCTACT 2689

RESULT 11
US-10-087-192-1231
; Sequence 1231, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1231
; LENGTH: 21410
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(21410)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-1231

Query Match 82.0%; Score 16.4; DB 13; Length 21410;
Best Local Similarity 94.4%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGCTTGTCTCAGAGCTACT 19
|||||
Db 9034 GGGCTTGTCTCAGAGCTACT 9051

RESULT 12
US-10-085-117-19
; Sequence 19, Application US/10085117
; Publication No. US2003023234A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 30486
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(30486)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-19

Query Match 82.0%; Score 16.4; DB 17; Length 30486;
Best Local Similarity 94.4%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 GGGCTTGTCTCAGAGCTACT 19

Db 6617 GGGCTTGTCTCAGAGCTACT 6634
|||||

RESULT 13
US-10-741-600-37582/c
; Sequence 37582, Application US/10741600
; Publication No. US20050026169A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001499
; CURRENT APPLICATION NUMBER: US/10/741,600
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 73997
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37582
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-741-600-37582

Query Match 79.0%; Score 15.8; DB 19; Length 201;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GGGCTTGTCTCAGAGCTACT 19
|||||
Db 24 GGGCTTGTCTCAGAGCTACT 6

RESULT 14
US-10-425-115-167738
; Sequence 167738, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 167738
; LENGTH: 213
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_84558C.1
US-10-425-115-167738

Query Match 79.0%; Score 15.8; DB 18; Length 213;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 GGGCTTGTCTCAGAGCTACT 20
|||||
Db 123 GGGCTTGTCTCAGAGCTACT 141

RESULT 15
US-09-732-627A-2660
; Sequence 2660, Application US/09732627A
; Publication No. US20040123338A1
; GENERAL INFORMATION:
; APPLICANT: Fincher, Karen L.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51770)B
US-09-732-627A-2660


```

; CURRENT APPLICATION NUMBER: US/09/732,627A
; CURRENT FILING DATE: 2000-12-08
; NUMBER OF SEQ ID NOS: 4930
; SEQ ID NO 2660
; LENGTH: 316
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(316)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3493-007-P1-M1-C5
US-09-732-627A-2660

```

```

Query Match      79.0%; Score 15.8; DB 11; Length 316;
Best Local Similarity 89.5%; Pred. No. 2.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy      1 CGGGCTTCTCAGAGCTACT 19
        |||||
Db      56 CGGGCTTCTCAGAGCTGCT 74

```

Search completed: March 13, 2005, 17:10:04
Job time : 15.0784 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 13, 2005, 03:11:12 ; Search time 97.2028 Seconds
(without alignments)
8399.558 Million cell updates/sec

Title: US-10-672-399-9

Perfect score: 20

Sequence: 1 cgggctgtcagagctactt 20

Scoring table: IDENTITY_NUC

Gapop 10_0 , Gapext 1.0

Searched: 4554873 seqs, 20411521753 residues

Total number of hits satisfying chosen parameters: 91109746

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_NA_Main:

- 1: /cgn2_6/ptodata/1/pna/PCTUS1 COMB.seq.*
- 2: /cgn2_6/ptodata/1/pna/PCTUS2 COMB.seq.*
- 3: /cgn2_6/ptodata/1/pna/PCTUS COMB.seq.*
- 4: /cgn2_6/ptodata/1/pna/US06 COMB.seq.*
- 5: /cgn2_6/ptodata/1/pna/US07 COMB.seq.*
- 6: /cgn2_6/ptodata/1/pna/US080 COMB.seq.*
- 7: /cgn2_6/ptodata/1/pna/US081 COMB.seq.*
- 8: /cgn2_6/ptodata/1/pna/US082 COMB.seq.*
- 9: /cgn2_6/ptodata/1/pna/US083 COMB.seq.*
- 10: /cgn2_6/ptodata/1/pna/US084 COMB.seq.*
- 11: /cgn2_6/ptodata/1/pna/US085 COMB.seq.*
- 12: /cgn2_6/ptodata/1/pna/US086 COMB.seq.*
- 13: /cgn2_6/ptodata/1/pna/US087 COMB.seq.*
- 14: /cgn2_6/ptodata/1/pna/US088 COMB.seq.*
- 15: /cgn2_6/ptodata/1/pna/US089 COMB.seq.*
- 16: /cgn2_6/ptodata/1/pna/US090 COMB.seq.*
- 17: /cgn2_6/ptodata/1/pna/US091 COMB.seq.*
- 18: /cgn2_6/ptodata/1/pna/US092A COMB.seq.*
- 19: /cgn2_6/ptodata/1/pna/US092B COMB.seq.*
- 20: /cgn2_6/ptodata/1/pna/US093A COMB.seq.*
- 21: /cgn2_6/ptodata/1/pna/US093B COMB.seq.*
- 22: /cgn2_6/ptodata/1/pna/US094 COMB.seq.*
- 23: /cgn2_6/ptodata/1/pna/US095A COMB.seq.*
- 24: /cgn2_6/ptodata/1/pna/US095B COMB.seq.*
- 25: /cgn2_6/ptodata/1/pna/US095C COMB.seq.*
- 26: /cgn2_6/ptodata/1/pna/US095D COMB.seq.*
- 27: /cgn2_6/ptodata/1/pna/US096A COMB.seq.*
- 28: /cgn2_6/ptodata/1/pna/US096B COMB.seq.*
- 29: /cgn2_6/ptodata/1/pna/US096C COMB.seq.*
- 30: /cgn2_6/ptodata/1/pna/US096D COMB.seq.*
- 31: /cgn2_6/ptodata/1/pna/US096E COMB.seq.*
- 32: /cgn2_6/ptodata/1/pna/US097A COMB.seq.*
- 33: /cgn2_6/ptodata/1/pna/US097B COMB.seq.*
- 34: /cgn2_6/ptodata/1/pna/US097C COMB.seq.*
- 35: /cgn2_6/ptodata/1/pna/US098A COMB.seq.*
- 36: /cgn2_6/ptodata/1/pna/US098B COMB.seq.*
- 37: /cgn2_6/ptodata/1/pna/US098C COMB.seq.*
- 38: /cgn2_6/ptodata/1/pna/US098D COMB.seq.*
- 39: /cgn2_6/ptodata/1/pna/US099A COMB.seq.*
- 40: /cgn2_6/ptodata/1/pna/US099B COMB.seq.*
- 41: /cgn2_6/ptodata/1/pna/US099C COMB.seq.*
- 42: /cgn2_6/ptodata/1/pna/US099D COMB.seq.*
- 43: /cgn2_6/ptodata/1/pna/US099E COMB.seq.*

- 44: /cgn2_6/ptodata/1/pna/US099F COMB.seq.*
- 45: /cgn2_6/ptodata/1/pna/US099G COMB.seq.*
- 46: /cgn2_6/ptodata/1/pna/US100A COMB.seq.*
- 47: /cgn2_6/ptodata/1/pna/US100B COMB.seq.*
- 48: /cgn2_6/ptodata/1/pna/US101A COMB.seq.*
- 49: /cgn2_6/ptodata/1/pna/US101B COMB.seq.*
- 50: /cgn2_6/ptodata/1/pna/US102A COMB.seq.*
- 51: /cgn2_6/ptodata/1/pna/US102B COMB.seq.*
- 52: /cgn2_6/ptodata/1/pna/US103A COMB.seq.*
- 53: /cgn2_6/ptodata/1/pna/US103B COMB.seq.*
- 54: /cgn2_6/ptodata/1/pna/US104A COMB.seq.*
- 55: /cgn2_6/ptodata/1/pna/US104B COMB.seq.*
- 56: /cgn2_6/ptodata/1/pna/US105A COMB.seq.*
- 57: /cgn2_6/ptodata/1/pna/US105B COMB.seq.*
- 58: /cgn2_6/ptodata/1/pna/US106A COMB.seq.*
- 59: /cgn2_6/ptodata/1/pna/US107A COMB.seq.*
- 60: /cgn2_6/ptodata/1/pna/US107B COMB.seq.*
- 61: /cgn2_6/ptodata/1/pna/US107C COMB.seq.*
- 62: /cgn2_6/ptodata/1/pna/US107D COMB.seq.*
- 63: /cgn2_6/ptodata/1/pna/US108A COMB.seq.*
- 64: /cgn2_6/ptodata/1/pna/US108B COMB.seq.*
- 65: /cgn2_6/ptodata/1/pna/US109A COMB.seq.*
- 66: /cgn2_6/ptodata/1/pna/US109B COMB.seq.*
- 67: /cgn2_6/ptodata/1/pna/US109C COMB.seq.*
- 68: /cgn2_6/ptodata/1/pna/US110 COMB.seq.*
- 69: /cgn2_6/ptodata/1/pna/US6000 COMB.seq.*
- 70: /cgn2_6/ptodata/1/pna/US6001 COMB.seq.*
- 71: /cgn2_6/ptodata/1/pna/US6002 COMB.seq.*
- 72: /cgn2_6/ptodata/1/pna/US6003 COMB.seq.*
- 73: /cgn2_6/ptodata/1/pna/US6004 COMB.seq.*
- 74: /cgn2_6/ptodata/1/pna/US6005 COMB.seq.*
- 75: /cgn2_6/ptodata/1/pna/US6006 COMB.seq.*
- 76: /cgn2_6/ptodata/1/pna/US6007 COMB.seq.*
- 77: /cgn2_6/ptodata/1/pna/US6008 COMB.seq.*
- 78: /cgn2_6/ptodata/1/pna/US6009 COMB.seq.*
- 79: /cgn2_6/ptodata/1/pna/US6010 COMB.seq.*
- 80: /cgn2_6/ptodata/1/pna/US6011 COMB.seq.*
- 81: /cgn2_6/ptodata/1/pna/US6012 COMB.seq.*
- 82: /cgn2_6/ptodata/1/pna/US6013 COMB.seq.*
- 83: /cgn2_6/ptodata/1/pna/US6014 COMB.seq.*
- 84: /cgn2_6/ptodata/1/pna/US6015 COMB.seq.*
- 85: /cgn2_6/ptodata/1/pna/US6016 COMB.seq.*
- 86: /cgn2_6/ptodata/1/pna/US6017 COMB.seq.*
- 87: /cgn2_6/ptodata/1/pna/US6018 COMB.seq.*
- 88: /cgn2_6/ptodata/1/pna/US6019 COMB.seq.*
- 89: /cgn2_6/ptodata/1/pna/US6020 COMB.seq.*
- 90: /cgn2_6/ptodata/1/pna/US6021 COMB.seq.*
- 91: /cgn2_6/ptodata/1/pna/US6022 COMB.seq.*
- 92: /cgn2_6/ptodata/1/pna/US6023A COMB.seq.*
- 93: /cgn2_6/ptodata/1/pna/US6023B COMB.seq.*
- 94: /cgn2_6/ptodata/1/pna/US6024 COMB.seq.*
- 95: /cgn2_6/ptodata/1/pna/US6025 COMB.seq.*
- 96: /cgn2_6/ptodata/1/pna/US6026 COMB.seq.*
- 97: /cgn2_6/ptodata/1/pna/US6027 COMB.seq.*
- 98: /cgn2_6/ptodata/1/pna/US6028 COMB.seq.*
- 99: /cgn2_6/ptodata/1/pna/US6029 COMB.seq.*
- 100: /cgn2_6/ptodata/1/pna/US6030 COMB.seq.*
- 101: /cgn2_6/ptodata/1/pna/US6031 COMB.seq.*
- 102: /cgn2_6/ptodata/1/pna/US6032 COMB.seq.*
- 103: /cgn2_6/ptodata/1/pna/US6033 COMB.seq.*
- 104: /cgn2_6/ptodata/1/pna/US6034 COMB.seq.*
- 105: /cgn2_6/ptodata/1/pna/US6035 COMB.seq.*
- 106: /cgn2_6/ptodata/1/pna/US6036 COMB.seq.*
- 107: /cgn2_6/ptodata/1/pna/US6037 COMB.seq.*
- 108: /cgn2_6/ptodata/1/pna/US6038 COMB.seq.*
- 109: /cgn2_6/ptodata/1/pna/US6039 COMB.seq.*
- 110: /cgn2_6/ptodata/1/pna/US6040 COMB.seq.*
- 111: /cgn2_6/ptodata/1/pna/US6041 COMB.seq.*
- 112: /cgn2_6/ptodata/1/pna/US6042 COMB.seq.*
- 113: /cgn2_6/ptodata/1/pna/US6043 COMB.seq.*
- 114: /cgn2_6/ptodata/1/pna/US6044 COMB.seq.*
- 115: /cgn2_6/ptodata/1/pna/US6045 COMB.seq.*
- 116: /cgn2_6/ptodata/1/pna/US6046 COMB.seq.*

117: /cgn2_6/ptodata/1/pna/US6047_COMB.seq.*
 118: /cgn2_6/ptodata/1/pna/US6048_COMB.seq.*
 119: /cgn2_6/ptodata/1/pna/US6049_COMB.seq.*
 120: /cgn2_6/ptodata/1/pna/US6050_COMB.seq.*
 121: /cgn2_6/ptodata/1/pna/US6051_COMB.seq.*
 122: /cgn2_6/ptodata/1/pna/US6052_COMB.seq.*
 123: /cgn2_6/ptodata/1/pna/US6053_COMB.seq.*
 124: /cgn2_6/ptodata/1/pna/US6054_COMB.seq.*
 125: /cgn2_6/ptodata/1/pna/US6055_COMB.seq.*
 126: /cgn2_6/ptodata/1/pna/US6056_COMB.seq.*
 127: /cgn2_6/ptodata/1/pna/US6057_COMB.seq.*
 128: /cgn2_6/ptodata/1/pna/US6058_COMB.seq.*
 129: /cgn2_6/ptodata/1/pna/US6059_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	20	58	US-10-672-399-9	Sequence 9, Appli
2	20	100.0	201	67	US-10-990-328-92497	Sequence 92497, A
3	20	100.0	201	67	US-10-990-328-92508	Sequence 92508, A
4	20	100.0	201	67	US-10-990-328-410752	Sequence 410752, A
5	20	100.0	201	115	US-60-452-680-115499	Sequence 115499, A
6	20	100.0	201	115	US-60-453-050-82424	Sequence 82424, A
7	20	100.0	201	115	US-60-453-135-82424	Sequence 82424, A
8	20	100.0	201	116	US-60-466-412-82424	Sequence 82424, A
9	20	100.0	201	116	US-60-466-412-402528	Sequence 402528, A
10	20	100.0	490	18	US-09-235-076-27210	Sequence 27210, A
11	20	100.0	490	19	US-09-289-768-29320	Sequence 29320, A
12	20	100.0	490	20	US-09-332-782-27210	Sequence 27210, A
13	20	100.0	490	33	US-09-737-223-27210	Sequence 27210, A
14	20	100.0	490	39	US-09-918-995-27210	Sequence 27210, A
15	20	100.0	490	40	US-09-939-397-29320	Sequence 29320, A
16	20	100.0	601	66	US-10-940-774-201780	Sequence 201780, A
17	20	100.0	1065	58	US-10-672-399-7	Sequence 7, Appli
18	20	100.0	1071	58	US-10-672-399-3	Sequence 3, Appli
19	20	100.0	1083	58	US-10-672-399-5	Sequence 5, Appli
20	20	100.0	1737	1	PCT-US02-41225A-223	Sequence 223, App
21	20	100.0	1737	1	PCT-US02-41225A-224	Sequence 224, App
22	20	100.0	1737	58	US-10-672-399-1	Sequence 1, Appli
23	20	100.0	1737	95	US-60-257-537-223	Sequence 223, App
24	20	100.0	1737	95	US-60-257-537-224	Sequence 224, App
25	20	100.0	2087	1	PCT-US02-41225A-226	Sequence 226, App
26	20	100.0	2087	49	US-10-170-235-37298	Sequence 37298, A
27	20	100.0	2087	66	US-10-940-774-5659	Sequence 5659, App
28	20	100.0	2087	95	US-60-257-537-226	Sequence 226, App
29	20	100.0	2087	115	US-60-452-680-12253	Sequence 12253, A
30	20	100.0	2087	115	US-60-453-050-7646	Sequence 7646, App
31	20	100.0	2087	115	US-60-453-135-7646	Sequence 7646, App
32	20	100.0	2087	116	US-60-466-412-7646	Sequence 7646, App
33	20	100.0	2088	1	PCT-US02-41225A-222	Sequence 222, App
34	20	100.0	2088	52	US-10-342-164-1	Sequence 1, Appli
35	20	100.0	2088	95	US-60-257-537-222	Sequence 222, App
36	20	100.0	2089	67	US-10-990-328-7459	Sequence 7459, App
37	20	100.0	2098	97	US-60-278-358-16271	Sequence 16271, A
38	20	100.0	2099	67	US-10-990-328-7460	Sequence 7460, App
39	20	100.0	2112	102	US-60-324-185-23093	Sequence 23093, A
40	20	100.0	2116	46	US-10-042-523-1	Sequence 1, Appli
41	20	100.0	14862	66	US-10-940-774-17401	Sequence 17401, A
42	20	100.0	22862	116	US-60-466-412-87508	Sequence 87508, A
43	20	100.0	22877	67	US-10-990-328-96152	Sequence 96152, A
44	20	100.0	231222	1	PCT-US02-41225A-225	Sequence 225, App
45	20	100.0	231222	95	US-60-257-537-225	Sequence 225, App

ALIGNMENTS

```

RESULT 1
US-10-672-399-9
; Sequence 9, Application US/10672399
; GENERAL INFORMATION:
; APPLICANT: University of Alberta
; TITLE OF INVENTION: Cancer Monitoring and Therapeutics
; FILE REFERENCE: A894635US
; CURRENT APPLICATION NUMBER: US/10/672,399
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: US 60/472,401
; PRIOR FILING DATE: 2003-05-22
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-672-399-9

Query Match      100.0%; Score 20; DB 58; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CGGGCTTGTGACAGCTACTT 20
      |||||
Db      1 CGGGCTTGTGACAGCTACTT 20

RESULT 2
US-10-990-328-92497
; Sequence 92497, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 92497
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-990-328-92497

Query Match      100.0%; Score 20; DB 67; Length 201;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 CGGGCTTGTGACAGCTACTT 20
      |||||
Db      180 CGGGCTTGTGACAGCTACTT 199

RESULT 3
US-10-990-328-92508
; Sequence 92508, Application US/10990328
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; TITLE OF INVENTION: POLYMORPHISMS IN NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING HUMAN ENZYME PROTEINS, METHODS OF DETECTION AND
; FILE REFERENCE: CL001495
; CURRENT APPLICATION NUMBER: US/10/990,328
; CURRENT FILING DATE: 2004-11-17
; NUMBER OF SEQ ID NOS: 558824
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 92508
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens

```

RESULT 6
US-60-453-050-82424
; Sequence 82424, Application US/60453050
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: LUKE, May

```
Query Match      100.0%; Score 20; DB 116; Length 201;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 180 CGGGCTTGTCTCAGAGCTACTT 199

RESULT 9
US-60-466-412-402528
; Sequence 402528, Application US/60466412
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: IAKOUBOVA, Olga
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: MYOCARDIAL INFARCTION, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001466
; CURRENT APPLICATION NUMBER: US/60/466,412
; CURRENT FILING DATE: 2003-04-30
; NUMBER OF SEQ ID NOS: 429241
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 402528
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-60-466-412-402528

Query Match 100.0%; Score 20; DB 116; Length 201;
Best Local Similarity 100.0%; Pred. No. 8.5; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 180 CGGGCTTGTCTCAGAGCTACTT 199

RESULT 10
US-09-235-076-27210
; Sequence 27210, Application US/09235076
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/235,076
; CURRENT FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27210
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-235-076-27210

Query Match 100.0%; Score 20; DB 18; Length 490;
Best Local Similarity 100.0%; Pred. No. 9.8; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 335 CGGGCTTGTCTCAGAGCTACTT 354

RESULT 11
US-09-289-768-29320
; Sequence 29320, Application US/09289768
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-765
; CURRENT APPLICATION NUMBER: US/09/289,768
```

```
; CURRENT FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 39996
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29320
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-289-768-29320

Query Match 100.0%; Score 20; DB 19; Length 490;
Best Local Similarity 100.0%; Pred. No. 9.8; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 335 CGGGCTTGTCTCAGAGCTACTT 354

RESULT 12
US-09-332-782-27210
; Sequence 27210, Application US/09332782A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/332,782A
; CURRENT FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27210
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(490)
; OTHER INFORMATION: n = A,T,C or G
US-09-332-782-27210

Query Match 100.0%; Score 20; DB 20; Length 490;
Best Local Similarity 100.0%; Pred. No. 9.8; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0;

Qy 1 CGGGCTTGTCTCAGAGCTACTT 20
Db 335 CGGGCTTGTCTCAGAGCTACTT 354

RESULT 13
US-09-737-223-27210
; Sequence 27210, Application US/09737223
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS cDNA LIBRARIES
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/737,223
; CURRENT FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 09/332,782
; PRIOR FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27210
; LENGTH: 490
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```


THIS PAGE BLANK (USPTO)